

2024 스마트미디어 추계학술대회

2024 Autumn Academic Conference
of Smart Media

스마트미디어 리터러시와 사회적 인간: 소통, 상생, 혁신

Smart Media Literacy and Social Humanity:
Communication, Coexistence, and Innovation

일 시 | 2024. 10. 11.(금) ~ 10. 12.(토)

장 소 | 한양대학교 경영관

주 최 | 한국전자거래학회, 한국스마트미디어학회

주최 |



Smart Media
KOREAN INSTITUTE OF SMART MEDIA
한국스마트미디어학회



사단법인한국전자거래학회
Society for e-Business Studies



한양대학교 LINC+사업단
HANYANG UNIVERSITY

구두 발표 0-1

○ 시 간 2024년 10월 12일 (토) 13:00~14:20

○ 발 표 장 한양대학교 경영관 7층 703호

○ 좌 장 윤봉식 교수(남부대학교)

01-01 제 목 향상된 고객 인게이지먼트를 위한 실시간 API 대화 기반 검색 증강 생성
저 자 김경선(다이퀘스트), 정한민(한국과학기술정보연구원)

01-02 제 목 Graph Transformer Networks for Optimized 5G Network Resource
Management: A Conceptual Study
저 자 Minghao Xu, Han-Gue Jo(Kunsan National University)

01-03 제 목 Deep Learning based BLE LOS/NLOS Classification in Indoor
Environments
저 자 Shawana Jamil, So-Yeon Kim, Jae-Young Pyun(Chosun University)

01-04 제 목 디자인씽킹기반 SW융합 교과 설계
저 자 유소월(전북대학교), 유장웅(남부대학교)

01-05 제 목 A Study of Efficient Federated Learning for Large Language Models
저 자 Shivani Sanjay Kolekar, Kyungbaek Kim(Chonnam National University)

구두 발표 O-2

○ 시 간 2024년 10월 12일 (토) 13:00~14:20

○ 발 표 장 한양대학교 경영관 7층 704호

○ 좌 장 변재영 교수(조선대학교)

O2-01 제 목 ASD 조기 예측을 위한 연합학습 기반 시선 추적 이미지 분류 모델
저 자 신현호, 이재민, 전홍우(한국과학기술정보연구원)

O2-02 제 목 다중 양식 융합과 개선된 E-MTN을 활용한 비디오 폭력 탐지 성능 향상
저 자 나광호, 최경주(충북대학교)

O2-03 제 목 차량 문콕사고 UX개선을 위한 연구방향 탐색
저 자 윤경빈, 윤봉식(남부대학교)

O2-04 제 목 아이템 영역 마스크를 이용한 딥러닝 기반 패션 아이템의 무늬 패턴 분류
저 자 권영채, 최진영, 김인중(한동대학교)

O2-05 제 목 ViT-Based Deep Learning Model for Potato Leaf Disease Detection
저 자 Md Ilias Bappi, Kyungbeak Kim(Chonnam National University)

O2-06 제 목 Multi-Class Dental Image Segmentation Using ResNet-50 and U-Net
저 자 Afnan Ghafoor, Bumshik Lee(Chosun University)

구두 발표 0-3

○ 시 간 2024년 10월 12일 (토) 13:00~14:20

○ 발 표 장 한양대학교 경제금융대학 1층 101호

○ 좌 장 김양훈 교수(신한대학교)

03-01 제 목 크리에이터 산업의 지속 가능성: 위기와 기회 (MCN 기업을 중심으로)
저 자 권병민, 이지은(한양사이버대학교)

03-02 제 목 문화적 푸시 및 풀 요인에 따른 여행 동기가 목적지 선택 만족도에 미치는 영향
저 자 강지효, 허지연, 차경진(한양대학교)

03-03 제 목 모바일 쇼핑몰의 환기성 높은 이미지가 원격현존감에 미치는 영향: VVIQ의 조절효과를 중심으로
저 자 김협(호서대학교), 권혁준(순천향대학교), 차욱재(성균관대학교), 박지섭(한국기술교육대학교)

03-04 제 목 방위산업기술 유출 현황 및보안 대책 연구동향 분석
저 자 김성현, 임희영, 김자원(대진대학교)

03-05 제 목 법제도 개선을 위한 텍스트 마이닝과 퍼지델파이 기반 장애인 분쟁유형 도출
저 자 황지선, 윤태영, 강주영(아주대학교)

구두 발표 O-4

○ 시 간 2024년 10월 12일 (토) 13:00~14:20

○ 발 표 장 한양대학교 경제금융대학 1층 102호

○ 좌 장 정재윤 교수(경희대학교)

O4-01 제 목 시계열 DB를 이용한 풍력발전 예측 머신러닝 시스템 개발
저 자 강명구, 정솔, 김민식, 정재윤(경희대학교)

O4-02 제 목 LLM을 이용한 고장진단의 설명가능 AI 기술 개발
저 자 김민식, 최서현, 정재윤(경희대학교)

O4-03 제 목 자산관리헬 표준을 이용한 자율이동로봇 모니터링
저 자 신민중, 한영석, 정재윤(경희대학교)

O4-04 제 목 YOLO 및 OpenCV 기반 냉장고 재활용 공정의 실시간 모니터링
저 자 김태연, 조운서, 장태우(경기대학교)

O4-05 제 목 수거 폐가전제품 정보관리 효율화를 위한 AI 기반 이미지 인식
저 자 이성훈, 김소현, 장태우(경기대학교)

구두 발표 O-5

○ 시 간 2024년 10월 12일 (토) 13:00~14:20

○ 발 표 장 한양대학교 경제금융대학 5층 501호

○ 좌 장 김효수 교수(중앙대학교)

O5-01 제 목 생성형 AI 기반 법률 지원 챗봇: 금융, 재난 관리 및 의료 분야에서의 BERT 기반 데이터 처리 및 응용에 대한 연구
저 자 박태윤, 유현욱, 김무철(중앙대학교)

O5-02 제 목 그래프 임베딩 개선을 위한 랜덤 워크의 내재된 편향성 완화 연구
저 자 류예종, 김무철(중앙대학교)

O5-03 제 목 합성 데이터셋 생성 및 이미지 분류를 이용한 도심 침수 탐지
저 자 강윤제, 정용훈, 서상현(중앙대학교)

O5-04 제 목 메타버스 내의 디지털 어뷰징 행동 탐지를 위한 학습 데이터 세트 구축 연구
저 자 조용규, 정용훈, 서상현(중앙대학교)

O5-05 제 목 각성도 및 긍·부정도를 적용한 멀티모달 감정인식 모델
저 자 신유담, 김범수, 서상현(중앙대학교)

구두 발표 O-6

○ 시 간 2024년 10월 12일 (토) 14:30~15:50

○ 발 표 장 한양대학교 경영관 7층 703호

○ 좌 장 김경백 교수(전남대학교)

O6-01 제 목 Optimizing 2D Object Detection for Autonomous Driving using Image Augmentation : Deep Drive Dataset Analysis and Real-Time Android Implementation

저 자 ERCELİK UGUR(Chonnam National University)

O6-02 제 목 Kano 모델 및 PCSI Index를 활용한 OTT 스포츠 중계 품질 개선 방안 연구

저 자 박미연, 최은정, 이수철(상명대학교)

O6-03 제 목 고령자 대상 다차원 인지 훈련 시스템의 교수자 경험 분석

저 자 김태홍, 정영재, 장경미, 도준형(한국한의학연구원)

O6-04 제 목 심층 강화 학습(DRL)을 이용한 동적 환경에서의 데이터 캐시 최적화

저 자 권민구, 송민석(인하대학교)

O6-05 제 목 채용공고 데이터마이닝을 활용한 테크니컬 스킬 요구에 대한 분석 -제조업 인공지능 관련 일자리 중심으로-

저 자 윤고은, 이충권(계명대학교)

구두 발표 0-7

○ 시 간 2024년 10월 12일 (토) 14:30~15:50

○ 발 표 장 한양대학교 경영관 7층 704호

○ 좌 장 조한규 교수(군산대학교)

07-01 제 목 호텔 예약 플랫폼에서 평점 수렴 상황에서 소비자 의사 결정에 관한 실험 연구
저 자 오호희, 고준(전남대학교)

07-02 제 목 디지털전환 관련 IT인력에 관한 연구 - 동남아시아 4개국 채용공고 분석 -
저 자 이진백, 이충권(계명대학교)

07-03 제 목 유럽 디자인 보호규정 개정이 디지털 디자인 권리에 미칠 영향에 대한 고찰
저 자 백은경(KAIST)

07-04 제 목 단안 비디오를 이용한 야구 피칭 모션 재건
저 자 김지원, 유리(아주대학교)

07-05 제 목 자연사 생태정보를 활용한 VR콘텐츠 연구
저 자 박기덕(유한대학교, 국립중앙과학관)

구두 발표 0-8

○ 시 간	2024년 10월 12일 (토) 14:30~15:50
○ 발 표 장	한양대학교 경제금융대학 1층 101호
○ 좌 장	권혁준 교수(순천향대학교)

08-01 제 목 은행에서 발생하는 금융 보고서 데이터의 정보 중요도에 따른 자동
등급화를 위한 ALBERT 기반 연구

저 자 전예림, 최예지, 이지민, 장항배(중앙대학교)

08-02 제 목 국내 사이버 회복력 개념 정립 및 평가체계 방안 연구

저 자 박지혜, 이상태, 장항배(중앙대학교)

08-03 제 목 다크 패턴과 피싱: 기술 유출의 새로운 사이버 보안 위협 분석

저 자 이지현, 변재욱, 장항배(중앙대학교)

08-04 제 목 개인정보보호 연구동향 분석: 국내 KCI 등재 학술지를 중심으로(2019-2023)

저 자 문지원, 박지혜, 장항배(중앙대학교)

08-05 제 목 국내 공공기관 개인정보 유출 사례를 통한 PET 도입 방안 연구

저 자 이다빈, 배성윤, 이주노, 장항배(중앙대학교)

구두 발표 O-9

○ 시 간	2024년 10월 12일 (토) 14:30~15:50
○ 발 표 장	한양대학교 경제금융대학 1층 102호
○ 좌 장	황규선 교수(울산대학교)

09-01	제 목 의사결정나무의 성능 향상을 위한 유전알고리즘 기반 이산화 방법론 저 자 류슬기, 천기정, 이철우, 최성근, 전성범(동국대학교)
09-02	제 목 통합시퀀스를 활용한 시점 별 선호도 차별화 추천시스템 저 자 문상혁, 구자협, 이영훈(서울과학기술대학교)
09-03	제 목 소규모 및 대규모 고객 집합을 가진 VRPPD 문제 해결을 위한 Greedy 알고리즘 및 하이브리드 유전자 알고리즘의 비교 분석 저 자 차병현, 이현경, 김해중(경기대학교)
09-04	제 목 AI자율제조 환경에서의 SDF(Software Defined Factory) 구축 전략에 관한 연구 저 자 하태원, 최서윤, 강현민, 황규선(울산대학교)
09-05	제 목 자율운반로봇의 관리 고도화를 위한 정체 관리 KPI 발굴 저 자 황규선(울산대학교)
09-06	제 목 병렬 챔버를 가진 반도체 클러스터 장비의 lot 투입 스케줄링 문제 저 자 황성보, 노윤서, 이가영, 하가영, 한준희(부산대학교)

구두 발표 O-10

○ 시 간	2024년 10월 12일 (토) 14:30~15:50
○ 발 표 장	한양대학교 경제금융대학 5층 501호
○ 좌 장	김진성 교수(중앙대학교)

O10-01 제 목 스마트폰 지자기 센서를 활용한 마그네틱 악세서리 식별 연구
저 자 최재민, 김효수(중앙대학교)

O10-02 제 목 웨어러블 인터페이스를 이용한 RF 신호 기반 접촉 부위 추정 기법 연구
저 자 신수현, 김효수(중앙대학교)

O10-03 제 목 혼합 현실 기기를 활용한 현실과 가상 공간 간의 좌표 매핑 기술 연구
저 자 성락빈, 이대원(중앙대학교)

O10-04 제 목 음성 기반 상호작용을 위한 디지털 휴먼 시스템 연구
저 자 우성미, 이대원(중앙대학교)

O10-05 제 목 이벤트 데이터 기반 인터랙티브 스포츠 스트리밍 웹 서비스 개발
저 자 신동민, 이대원(중앙대학교)

구두 발표 O-11

○ 시 간	2024년 10월 12일 (토) 16:00~17:20
○ 발 표 장	한양대학교 경영관 7층 703호
○ 좌 장	나인호 교수(군산대학교)

O11-01 제 목 AI를 이용한 게이미피케이션 교수학습법 연구 모형 설계
저 자 윤봉식(남부대학교), 김진경(목포대학교)

O11-02 제 목 중국 크로스보더 전자상거래 플랫폼이 소비자의 구매의도에 미치는 영향 연구
 -중국 온라인 플랫폼B2C형을 중심으로-
저 자 진티엔러, 박복재(전남대학교)

O11-03 제 목 한국 전래동화 소재의 보드게임 개발을 위한 판타지적 특징 분석
저 자 이유안, 한지애(조선대학교)

O11-04 제 목 게이미피케이션 교육콘텐츠에서 교육내용-전략-게임요소 관계
저 자 한지애(조선대학교)

O11-05 제 목 특허정보를 활용한 국내 반려동물 산업 동향에 관한 연구:
 네트워크 분석과 텍스트 마이닝 기법을 활용하여
저 자 이동현, 강주영, 이한솔(아주대학교)

구두 발표 O-12

○ 시 간 2024년 10월 12일 (토) 16:00~17:20

○ 발 표 장 한양대학교 경영관 7층 704호

○ 좌 장 장준혁 교수(한남대학교)

O12-01 제 목 글로벌 전자상거래가 중국의 수출에 미치는 영향
저 자 하작호, 박복재(전남대학교)

O12-02 제 목 KOC(키 오피니언 소비자)특성이 구매의도에 미치는 영향 및 소비자 성별의
조절효과 -샤옹수를 중심으로-
저 자 저위성, 박복재(전남대학교)

O12-03 제 목 은유적 내러티브를 활용한 ESG 환경 영상 디자인
저 자 문성예, 한지애(조선대학교)

O12-04 제 목 산업환경에서 개인 보호 장비 감지를 위한 개선된 YOLOv8 모델 설계
저 자 박석우, 이영석, 나인호(군산대학교)

O12-05 제 목 초기 및 후기 융합 기법을 활용한 개선된 MDVC 모델
저 자 조성민, 최경주(충북대학교)

구두 발표 O-13

○ 시 간 2024년 10월 12일 (토) 16:00~17:20

○ 발 표 장 한양대학교 경제금융대학 1층 101호

○ 좌 장 김협 교수(호서대학교)

O13-01 제 목 클라우드 관점에서의 스마트리터러시에 관한 역량설계 연구
저 자 장계훈(중앙대학교)

O13-02 제 목 효율적인 해킹공격 대응을 위한 LLM 기반 허니팟 구축 아키텍처 연구
저 자 이진욱(중앙대학교)

O13-03 제 목 보안관제 시스템에서 방화벽 로그 분석을 위한 AI 적용 방안
저 자 김지훈(중앙대학교)

O13-04 제 목 제1금융권과의 비교를 통한 제2금융권의 적정 보안 체계 수립 및 실무적
보안 강화 방안: 중소 금융기관을 위한 효과적 대응 전략
저 자 이은지(중앙대학교)

O13-05 제 목 호텔 내부자의 악의적 행위 리스크 평가: 심리적 계약 위반 이론 중심으로
저 자 권종훈(중앙대학교)

O13-06 제 목 AI 기술을 이용한 보안인증제도 활성화 방안
저 자 안진우(중앙대학교)

구두 발표 O-14

○ 시간	2024년 10월 12일 (토) 16:00~17:20
○ 발표장	한양대학교 경제금융대학 1층 102호
○ 좌장	박지섭 교수(한국기술교육대학교)

- | | |
|--------|--|
| O14-01 | 제 목 불만에서 혁신으로:소비자의 부정적 리뷰 분석 기반 전기차 시장 확대 방안
저 자 왕문준, 최재원(순천향대학교) |
| O14-02 | 제 목 서사적 매력과 광고 전파력: 소셜 미디어에서의 서사 수준이 바이럴 효과에 미치는 영향
저 자 유몽, 최재원(순천향대학교) |
| O14-03 | 제 목 파트너십 구축: 공동 브랜딩에서 성공을 이끄는 요인
저 자 진연여, 최재원(순천향대학교) |
| O14-04 | 제 목 사용자를 위한 ASMR 마케팅 콘텐츠의 에코챔버 효과: 텍스트 마이닝 관점에서의 분석
저 자 르키아, 최재원(순천향대학교) |
| O14-05 | 제 목 도덕 결여 행동과 브랜드 위기: 소셜 미디어 인플루언서 스캔들이 소비자 신뢰에 미치는 영향
저 자 이맹릉, 최재원(순천향대학교) |

구두 발표 O-15

○ 시 간 2024년 10월 12일 (토) 16:00~17:20

○ 발 표 장 한양대학교 경제금융대학 5층 501호

○ 좌 장 이대원 교수(중앙대학교)

O15-01 제 목 서버리스 컴퓨팅에서 콜드 스타트 문제를 다루기 위한 기법들에 대한 분석
저 자 조민규, 박상오(중앙대학교)

O15-02 제 목 무인 항공기의 협력적 통신에 관한 연구
저 자 김동현, 박상오(중앙대학교)

O15-03 제 목 디지털 트윈을 활용한 고부가가치 저차원 소재의 물성 예측
저 자 이재원, 양승화(중앙대학교)

O15-04 제 목 인실리코 해석을 이용한 방열복합재료의 물성 예측
저 자 기나현, 양승화(중앙대학교)

포스터 발표 P-1

○ 시 간	2024년 10월 12일 (토) 14:30~15:50
○ 발 표 장	한양대학교 경영관 7층 SKT Hallway
○ 좌 장	김영철 교수(홍익대학교)

P1-01	제 목 2024년 소비 트렌드를 반영한 포스트 휴먼 시대에 AI를 활용한 과학기술 정책전략 수립 및 디지털 대전환 방안에 관한 연구 저 자 최장원, 황명권, 최선희(한국과학기술정보연구원)
P1-02	제 목 기억재생 나노칩 R&D 로드맵 수립 및 추진전략: 정책적 제언 저 자 최장원, 황명권, 최선희(한국과학기술정보연구원)
P1-03	제 목 인공지능 윤리 글로벌 공동연구 네트워크 탐색 저 자 유화선, 광영, 정도범(한국과학기술정보연구원)
P1-04	제 목 KDD/KM 방법론을 활용한 미래차 분야 국제협력 연구동향 분석 저 자 유화선(한국과학기술정보연구원)
P1-05	제 목 정밀 휴먼 디지털 트윈 구현을 위한 비전 트랜스포머 기반 의료영상 초해상화 모델: 무릎 영상 사례 저 자 이재용, 양동현, 심형섭, 황명권, 이경하(한국과학기술정보연구원)
P1-06	제 목 페르소나를 이용한 인공지능 문헌 검토자의 적절성 연구 저 자 이태석, 선충녕, 한상준, 유수현(한국과학기술정보연구원)
P1-07	제 목 흐린 이미지 재구성을 위한 3D 가우시안 스플래팅 향상 저 자 원일남, 조인휘(한양대학교)
P1-08	제 목 스타일 전환이 가능한 3D 객체 재구성 저 자 권순양, 조인휘(한양대학교)
P1-09	제 목 LDA 모델 기반 디지털 트윈 연구 트렌드 분석 저 자 고미현(한국과학기술정보연구원), 이수철(상명대학교)
P1-10	제 목 동적 벡터 색인을 이용한 지능형 학술검색 저 자 선충녕, 이태석, 한상준, 유수현(한국과학기술정보연구원)

포스터 발표 P-1

○ 시 간	2024년 10월 12일 (토) 14:30~15:50
○ 발 표 장	한양대학교 경영관 7층 SKT Hallway
○ 좌 장	김영철 교수(홍익대학교)

P1-11	제 목 YOLOv8-pose 모델의 P6 구조 및 SimAM 모듈 기반 성능 개선 저 자 조상호, 조인휘(한양대학교)
P1-12	제 목 인공지능 개발 지원을 위한 통합 플랫폼 운영 전략 저 자 황명권, 최장원, 최선희(한국과학기술정보연구원)
P1-13	제 목 멀티모달 프롬프트 학습에 대한 재고하기 비전 언어 모델 저 자 저우형, 김상철(전북대학교)
P1-14	제 목 중소기업 기업가 지향성이 비즈니스모델 혁신과 기업가적 성과에 미치는 영향: 조직문화 조절효과 저 자 곽영, 유화선, 정도범(한국과학기술정보연구원)
P1-15	제 목 Robust Path Tracking and Control for Autonomous Four-Wheel Irrigation Robots in a Gazebo Simulation Environment 저 자 Muhammad Ibrahim Zain Ul Abideen, SangCheol Kim(Jeonbuk National University)
P1-16	제 목 Research on Improving Intelligent Terminal Service Design Based on Context Awareness to Enhance the Lives of Older Adults -Focusing on Intelligent Terminals Used in Nursing Homes- 저 자 Zi-An Zhu, Byung-Pyo Kyung(Kongju National University)
P1-17	제 목 축사 환경에서 LSTM 오토인코더 기반 노이즈 감지와 오디오 인페인팅 저 자 정세연, 김상철(전북대학교)
P1-18	제 목 Deep Learning for Pig Face Recognition: A Comprehensive Detection and Classification Approach 저 자 Ruihan Ma, Sang-Cheol Kim(Jeonbuk National University)
P1-19	제 목 생성형 인공지능 기반 기술발전단계 평가 연구 저 자 이하윤, 이주현, 양혜영(한국과학기술정보연구원)

포스터 발표 P-2

○ 시 간	2024년 10월 12일 (토) 16:00~17:20
○ 발 표 장	한양대학교 경영관 7층 SKT Hallway
○ 좌 장	백은경 교수(KAIST)

P2-01	제 목 군집화와 객체 인식을 이용한 전방위 융합 객체 인식 시스템 저 자 김동훈, 조세운, 김병준, 김서정, 조연호, 현상미, 배성재(한국전자기술연구원), 최경호, 최나영 (에코스ENC), 정성환(한국전자기술연구원)
P2-02	제 목 Autonomous Navigation System Using Dual Extended Kalman Filter with RTK GNSS 저 자 Hassan Ali, SangCheol Kim(Jeonbuk National University)
P2-03	제 목 이차전지 산업의 최상위 수출입 네트워크 분석 저 자 박 완, 전홍우, 권이남(한국과학기술정보연구원)
P2-04	제 목 웹기반 Google OTP 관리 시스템 구축 저 자 김애숙(한남대학교)
P2-05	제 목 응급환자 이송을 위한 해양원격의료지도시스템 실증 사례 분석 저 자 박건우, 최종명, 김선경, 이영호(목포대학교)
P2-06	제 목 Automated Shrimp Size Measurement Using Semantic Graphics in Real Time 저 자 Waqar Malik Muhammad, SangCheol Kim(Jeonbuk National University)
P2-07	제 목 Study on the Effectiveness of Retrieval-Augmented Generation 저 자 Athiruj Poositaporn, Hanmin Jung, Jung Hoon Park(KIST)
P2-08	제 목 과학기술 이슈대응 정보서비스의 이용 현황 분석 저 자 윤정선, 배수영, 한상준, 유수현(한국과학기술정보연구원)
P2-09	제 목 효율적 조류 객체 탐지를 위한 YOLO 모델 비교 연구 저 자 왕흥요, 이승재, 유지형, 유경호, 김형주, 김판구(조선대학교)
P2-10	제 목 MLCA 기법을 적용한 향상된 조류 객체 탐지모델 연구 저 자 왕흥요, 이승재, 유지형, 유경호, 김형주, 김판구(조선대학교)

포스터 발표 P-2

○ 시 간 2024년 10월 12일 (토) 16:00~17:20

○ 발 표 장 한양대학교 경영관 7층 SKT Hallway

○ 좌 장 백은경 교수(KAIST)

P2-11 제 목 건설 XR 협업 서비스를 위한 웹 포털시스템

저 자 서명배, 박형진(한국건설기술연구원)

P2-12 제 목 머신러닝을 활용한 흙의 알파산 함량 예측 모델 개발에 관한 연구

저 자 이현조(한국농수산대학교), 백하영(강원대학교),
주원균(한국과학기술정보연구원), 구현정, 채철주(한국농수산대학교)

P2-13 제 목 머신 러닝 기반 넙치 질병 분류 알고리즘

저 자 정종호, 김찬진(빛가람정보㈜), 손현승, 임한규(목포대학교),
최한석(빛가람정보㈜)

포스터 발표 P-3

○ 시 간	2024년 10월 12일 (토) 16:00~17:20
○ 발 표 장	한양대학교 경제금융대학 5층 복도
○ 좌 장	서상현 교수(중앙대학교)

P3-01	제 목 인공지능의 활용이 세금신고업무에 미치는 영향 저 자 장성일(단돈세무회계), 김정연(상명대학교)
P3-02	제 목 CUDA Core와 Tensor Core를 사용한 두 개의 소규모 행렬의 곱 연산 성능 비교 저 자 김인서, 김진성(중앙대학교)
P3-03	제 목 혼합 정밀도 및 양자화를 활용한 객체 탐지 모델 성능 비교 연구 저 자 박준형, 김진성(중앙대학교)
P3-04	제 목 머신러닝을 활용한 새만금 지역의 압축지수 예측 저 자 류승연(중앙대학교)
P3-05	제 목 유한요소 해석을 이용한 말뚝의 주면마찰력에 영향을 미치는 매개변수 연구 저 자 이태겸(중앙대학교)
P3-06	제 목 제스처로 유발된 표면의 진동을 기반으로 한 허공 제스처 인식 연구 저 자 김민혁, 김효수(중앙대학교)
P3-07	제 목 대형 언어 모델에서 가지치기와 혼합 정밀도 기법을 활용한 성능 비교 연구 저 자 박준형, 김진성(중앙대학교)
P3-08	제 목 하드웨어와 소프트웨어 관점에서 본 GPU 에너지 소비량 감축 방법에 대한 연구 저 자 김인서, 김진성(중앙대학교)
P3-09	제 목 합성데이터 기반 소나무재선충병 탐지 저 자 정용훈, 서상현(중앙대학교)
P3-10	제 목 델타 맵 기반 고충실도 비디오 얼굴 재노화 저 자 김범수, 서상현(중앙대학교)

포스터 발표 P-3

○ 시 간 2024년 10월 12일 (토) 16:00~17:20

○ 발 표 장 한양대학교 경제금융대학 5층 복도

○ 좌 장 서상현 교수(중앙대학교)

P3-11 제 목 합성데이터를 활용한 실내 시설 내 방문자 탐지
저 자 신원섭, 강윤제, 서상현(중앙대학교)

P3-12 제 목 서버리스 컴퓨팅을 위한 워크로드 예측 기술 동향 조사
저 자 조민규(중앙대학교), 이재환(국립공주대학교), 박상오(중앙대학교)

P3-13 제 목 지능형 문서처리 플랫폼의 중요정보 보호 방안 연구
저 자 홍기완(중앙대학교), 한소영(중앙대학교), 김자원(대진대학교),
김양훈(신한대학교), 장항배(중앙대학교)

학부생 구두발표 및 동영상 발표

○ 시간	2024년 10월 12일 (토) 16:00~17:20
○ 발표장	한양대학교 경영관 7층 707호
○ 좌장	차경진 교수(한양대학교)

O-01	제 목 탐색적 데이터 분석(EDA)기반 흡연상태와 음주여부에 따른 BMI수치 분석 저 자 전석원, 장준혁(한남대학교)
O-02	제 목 비지도 기계학습을 이용한 뇌 영상 기반 자폐 스펙트럼 장애의 생물학적 특성 분석기술 개발 저 자 최충민, 김상우, 김예은, 조한규(군산대학교)
O-03	제 목 수면 지표를 활용한 신체활동 예측: 스마트워치를 활용한 탐색적 연구 저 자 정현우, 손상진, 김예은, 이지원, 조한규(군산대학교)
O-04	제 목 LLaMA3 모델 기반의 Private LLM 시스템 설계 저 자 정도윤(호남대학교), 이형호(㈜가민정보시스템), 정희자(㈜휴넷가이아), 김남호(호남대학교)
O-05	제 목 YOLO를 통한 이미지 분석 기반 유실물 탐지 시스템 설계 저 자 조민성, 김민재, 김진성, 박정수, 김민서, 민홍(가천대학교)
O-06	제 목 자폐스펙트럼 장애 진단 그래프 신경망 모델 비교 연구 저 자 김다연, 백빈나, 배하람, 유채린, 이지원, 조한규(군산대학교)
O-07	제 목 디지털 트윈 도시에서 3D 객체와 실시간 스트림처리의 의미론적 통합 저 자 이예원, 이성범, 한정규, 천세진(동아대학교)
O-08	제 목 정보 취약 계층을 위한 블록체인 기반 노인 구직 시스템 설계 저 자 조민성, 김민서, 김민재, 이영찬, 민홍(가천대학교)
O-09	제 목 인간 정보처리 특성을 고려한 안전 정보디자인: 감각, 지각, 인지, 기억단계별 디자인 전략 저 자 김다솔, 류시천(조선대학교)

포스터 발표 U-1: 학부생 논문

○ 시간	2024년 10월 12일 (토) 13:00~14:20
○ 발표 장	한양대학교 경영관 7층 SKT Hallway
○ 좌 장	최재혁 교수(조선대학교)

U1-01	제 목 신입생 및 복학생의 학교 적응을 위한 소모임 어플리케이션 저 자 이규민, 김종윤, 이민우, 심훈, 전태민, 송지영(한남대학교)
U1-02	제 목 안면 색상 데이터 기반 퍼스널 컬러 진단 서비스 저 자 김주희, 김채하, 이가은, 송지영(한남대학교)
U1-03	제 목 초음파 센서 기반 실시간 빗물받이 관리 시스템 저 자 배정훈, 최지영, 최지민, 국지언, 장민성, 송지영(한남대학교)
U1-04	제 목 클라우드 백엔드 비용 최적화를 위한 클라이언트 캐싱 기법 설계 및 구현 저 자 서지웅, 장준혁(한남대학교)
U1-05	제 목 저학년 코딩교육을 위한 가상현실 기반 학습 방법 연구 저 자 홍서연, 박현아, 최지영, 최미서, 김장환, 서채연, 김영철(홍익대학교)
U1-06	제 목 Video Swin Transformer를 이용한 올바른 운동 자세 추정 저 자 윤혜성, 백으뜸(호남대학교)
U1-07	제 목 로컬 어텐션 기반의 밀집 해충 카운팅 모델 저 자 조정민, 김경태, 최원준, 손창환(군산대학교)
U1-08	제 목 요양시설 내 노약자 낙상 감지 모니터링 시스템 저 자 박준렬, 박민현, 이민혁, 전민서, 양진모, 김영철(홍익대학교)
U1-09	제 목 청각장애 아동을 위한 수화 교육용 온라인 프로그램 개발 저 자 강성연, 김서영, 정현정, 이은서, 송지영(한남대학교)
U1-10	제 목 반려견 감정 인식을 위한 CNN-LSTM 기반 이미지 캡션화 기술 개발 저 자 오승기, 백으뜸(호남대학교)

포스터 발표 U-1: 학부생 논문

○ 시간	2024년 10월 12일 (토) 13:00~14:20
○ 발표장	한양대학교 경영관 7층 SKT Hallway
○ 좌장	최재혁 교수(조선대학교)

U1-11	제 목 맞춤형 디지털 리터러시를 위한 LLM과 지식 그래프를 활용한 프레임워크 설계 저 자 임가겸, 박정현, 한정규, 천세진(동아대학교)
U1-12	제 목 OTT 기반의 모션 검출 및 객체인식 모델 연구 저 자 김유림, 이승재, 최휘녕(전북대학교)
U1-13	제 목 YOLOv8 기반 실시간 태양광 패널 오염물질 인식 및 자동 청소 로봇 저 자 송상진, 민석주(호남대학교), 정희자(㈜휴넷가이아), 최광미, 김남호(호남대학교)
U1-14	제 목 자연어 기반 맞춤형 도서 추천을 위한 생성형AI 사서 저 자 류동훈, 박소윤, 김성은, 황두현(홍익대학교), 공지훈(투스퀘어), 김기두(한국정보통신 기술협회), 김장환, 서채연, 김영철(홍익대학교)
U1-15	제 목 전통문화 체험을 위한 실감형 콘텐츠 디지털 헤리티지 적용사례 저 자 우재영, 한진호, 위종우, 이준원, 김주연(홍익대학교), 공지훈(투스퀘어), 김장환, 서채연, 김영철(홍익대학교)
U1-16	제 목 가상 공유 갤러리 상에서 테마와 태그 기반 효율적 이미지 검색 저 자 한진우, 박민혁, 홍승빈, 우하은, 공지훈, 김장환, 서채연, 김영철(홍익대학교)
U1-17	제 목 전기차량에 적용된 케블라 섬유 열선 히터 불량 검출기 개발 및 제작 저 자 최우성(군산대학교), 박재완(㈜재현이노텍), 이준희, 성명훈, 김재원, 조소현, 우 주, 정재훈(군산대학교)
U1-18	제 목 애장품 소유 감성 메커니즘을 적용한 게임 제작 사례 저 자 손희완, 은광하(상명대학교)
U1-19	제 목 인터랙티브 드라마 <Deleter> 개발사례 분석 저 자 김도혜, 조성주, 허영(상명대학교)

포스터 발표 U-2: 학부생 논문

○ 시간	2024년 10월 12일 (토) 13:00~14:20
○ 발표장	한양대학교 경제금융대학 5층 복도
○ 좌장	박형일 팀장(KT)

U2-01	제 목 YOLOv8 기반 배전공기 관리 시스템 저 자 이훈제, 조성민, 정세훈, 심춘보(순천대학교)
U2-02	제 목 CoD Prompt와 Unsloth를 활용한 미세 조정된 Gemma2 기반의 한국어 문서 요약 효율화 저 자 김승주, 정세훈, 심춘보(순천대학교)
U2-03	제 목 동화 대탈출: 제스처 인식을 활용한 아동 단어 학습 게임 저 자 조하진, 김리연, 엄성용(서울여자대학교)
U2-04	제 목 자율주행 상황 속 환경 객체와 상호작용하는 모빌리티 AR 콘텐츠 개발 방법 연구 저 자 박유천, 경병표(공주대학교)
U2-05	제 목 ChatGPT를 활용한 인공지능 건강관리 애플리케이션 설계 저 자 이호연, 한선아, 정민교(서울여자대학교)
U2-06	제 목 생성형 AI를 활용한 게임 시나리오 생성에서 사전 학습과 계층적 프롬프팅이 사용자 경험에 미치는 영향 저 자 허주희, 이진주, 이소정, 정민교(서울여자대학교)
U2-07	제 목 생성형 AI 기반의 시험문제와 핵심 요약 제작 서비스 저 자 이채영, 김민영, 정민교(서울여자대학교)
U2-08	제 목 정보화 솔루션을 사용하는 생산현장 내 근무특성을 고려한 음성과 모션 Data 기술 저 자 장정우, 하호수, 임해나, 조훈희(전북대학교), 김용현, 김예진(㈜블루비즈)
U2-09	제 목 제품이미지를 활용한 웹드라마 광고콘텐츠 제작사례수 저 자 이채이, 이채영, 은광하(상명대학교)
U2-10	제 목 프로젝션 맵핑 기반 고전 게임의 표현 적용을 위한 시각표현 연구 저 자 김동민, 박수민, 이규서, 은광하(상명대학교)

포스터 발표 U-2: 학부생 논문

○ 시간	2024년 10월 12일 (토) 13:00~14:20
○ 발표장	한양대학교 경제금융대학 5층 복도
○ 좌장	박형일 팀장(KT)

U2-11	제 목 시각장애인의 게임 접근성 향상을 위한 핸드트래킹 게임 개발 저 자 곽수연, 김현진, 김연아, 황준(서울여자대학교), 고진광(순천대학교)
U2-12	제 목 졸업 프로젝트 아카이빙과 신입 인재 채용을 위한 양방향 매칭 플랫폼에 대한 연구 저 자 김주은, 김희주, 황준(서울여자대학교)
U2-13	제 목 AI 기반 반려묘 피부 증상 탐지 서비스 저 자 최소원, 박유진, 이예림, 엄성용(서울여자대학교)
U2-14	제 목 유학생을 위한 문화 교류 매칭 및 커뮤니티 서비스 구현 저 자 윤부영, 강예림, 김수연, 황준(서울여자대학교)
U2-15	제 목 캡 : 이주 외국인교과 선생님을 위한 한국어 숙제 지원 애플리케이션 저 자 이윤지, 석민솔, 신지우, 박시영, 엄성용(서울여자대학교)
U2-16	제 목 디지털트윈 기반 파이프랙 강구조물의 동특성 모니터링 시스템 구축 저 자 송민호, 이서진, 양재영, 김현우(조선대학교), 박재용(스마트오션), 최재혁(조선대학교)
U2-17	제 목 디지털트윈 레벨3 구축을 위한 s-BIM 모델 생성과 실시간 모니터링 저 자 권진영, 허유정, 장현석, 김용주, 양재영, 최재혁(조선대학교)

포스터 발표 U-3: 학부생 논문

○ 시간	2024년 10월 12일 (토) 14:30~15:50
○ 발표장	한양대학교 경제금융대학 5층 복도
○ 좌장	김무철 교수(중앙대학교)

U3-01	제 목 기기 에이전트 - 게이트웨이 배치 모델을 활용한 차등 프라이버시 적용 방안 연구 저 자 박시형(중앙대학교)
U3-02	제 목 대출가망 예측을 위한 랜덤포레스트와 차등 프라이버시의 융합:안전한 데이터 활용 방안 저 자 이지용(중앙대학교)
U3-03	제 목 K-익명성 기반의 차분 프라이버시를 이용한 안전한 개인정보 활용 저 자 장승수(중앙대학교)
U3-04	제 목 GAN 기반 데이터 증강을 활용한 개인정보 보호 강화 및 멤버십 추론 공격 방어 기술 연구 저 자 전하영(중앙대학교)
U3-05	제 목 차등 프라이버시와 데이터 합성 기술을 적용한 신경망 모델의 성능 및 프라이버시 보호 분석 저 자 조성원(중앙대학교)
U3-06	제 목 딥페이크 범죄 경로 분석을 통한 다차원적 대응 방안 제언 저 자 최강규(중앙대학교)
U3-07	제 목 여행조각가: LLM 기반 여행 영상 스토리보드 제작 서비스 저 자 김도균, 조민규, 박상오(중앙대학교)
U3-08	제 목 Char각: 머신러닝 기반 촬영 가이드 플랫폼 저 자 정현서, 김경연, 김동현, 박상오(중앙대학교)
U3-09	제 목 게이트웨이: 웹2와 웹3의 연결 저 자 추교현, 조민규, 박상오(중앙대학교)

향상된 고객 인게이지먼트를 위한 실시간 API 대화 기반 검색 증강 생성

김경선, 정한민
다이퀘스트, 한국과학기술정보연구원
e-mail : kksun@diquest.com, jhm@kisti.re.kr

Retrieval-Augmented Generation Based on Real-Time API Communication for Enhanced Client Engagement

Kyungsun Kim, Hanmin Jung
DiQuest Inc., Korea Institute of Science and Technology Information

요약

본 연구는 대규모 데이터를 보유한 서비스와 검색 증강 생성(RAG) 간의 효율적이고 확장 가능한 연계를 제안한다. 기존 RAG의 한계를 극복하기 위해, 외부 데이터를 실시간으로 API를 통해 연계하여 처리함으로써 자원 관리의 효율성, 데이터 최신성 유지, 보안 강화, 시스템 확장성 등의 이슈를 해결하고자 한다. 특히, 자체 구축한 RAG와 한국과학기술정보연구원의 ScienceON API Gateway를 연계하여 고객 인게이지먼트 전략 보고서 생성 과정과 결과를 보임으로써 본 제안의 효용성을 입증하였다.

1. 서론

ChatGPT가 촉발한 LLM(Large Language Model) 연구와 활용이 급격히 증가하면서, 환각(Hallucination)이나 과거 데이터로 인한 결과의 부정확성이 이슈로 대두되고 있다[1]. 이를 해결하기 위해 검색을 통하여 확보한 현실 지식을 이용하여 LLM을 증강시키는 방식으로 검색 증강 생성(RAG; Retrieval-Augmented Generation)이 제안되어, 최근 활발히 연구되고 있다[2-3].

그렇지만, 이론적으로 외부 정보 서비스로부터 데이터를 확보하여 내부 데이터와 통합하는 방식의 검색 증강 생성은 다음과 같은 한계를 가질 수밖에 없다. 먼저, 외부 정보 서비스가 활용하는 데이터를 보안이나 공개 측면에서 모두 제공받는 게 쉽지 않다. 예를 들어, 한국과학기술정보연구원의 ScienceON은 1.3억 건의 국내·외 논문과 4천만 건의 국내·외 특허 등을 보유하고 있으나¹⁾, 이를 모두 외부 기관에 제공하는 것은 저작권 이슈 등으로 불가능하다. 또한, 이러한 방대한 양의 데이터를 제공받는다고 하더라도 내부 데이터와의 통합, 학습, 분석은 엄청난 컴퓨팅 자원을 요구하기 때문에 시스템으로 구현하는 것도 쉽지 않다. 전자도서관, 특히 검색 시스템, 의료 데이터베이스 등과 같은 대규모 데이터를 보유한 분야들에서 공통으로 겪을 수밖에 없는 이슈이며, 이를 해결할 새로운 검색 증강 생성 방법이 필요하다.

이에 본 연구는 대규모 데이터를 보유한 서비스와 RAG 간의 효율적이고 확장 가능한 연계를 제안하는 것을 목표로 한다. 특히, 데이터를 직접 통합하지 않고, 필요 시점에 API를 통해 실시간으로 데이터를 요청하는 방법을 제안함으로써, 자원 관리의 효율성을 극대화하고 데이터의 최신성을 유지하며, 보안 요구사항을 충족시키고자 한다.

특히, 고객 인게이지먼트(Client Engagement)²⁾ 전략 보고서 생성을 통해 효용성을 보임으로써 본 연구의 우수성을 증명하고자 한다.

2. 실시간 API 대화 기반 검색 증강 생성

2.1 제안 필요성

1. 자원 관리 효율성 향상: 대규모 데이터를 RAG에 직접 통합할 경우, 데이터 저장 및 처리 과정에서 막대한 자원 소비가 발생할 수 있으며, 이에 따라 시스템 운영 비용이 증가하고, 처리 시간이 길어져 실시간 응답 성능이 저하될 수 있다. API를 통한 실시간 데이터 요청 방식은 이러한 문제를 해결할 수 있는 효과적인 방법으로, 필요한 시점에 필요한 데이터만을 가져와 처리함으로써 자원의 효율적 사용을 가능하게 한다.

2. 데이터 최신성 유지: 데이터를 RAG에 직접 통합하는 방식은 외부에서 생산된 데이터를 반입하고 갱신하는 주기에 따라 데이터의 최신성을 보장하는 데 한계가 있다. API를 통한 실시간 데이터 요청 방식은 항상 최신의 데이터를 사용하여 답변을 생성할 수 있어, 정보의 유효성과 신뢰성을 높일 수 있다.

3. 보안 및 개인정보 보호: 민감한 데이터를 포함하는 의료, 금융 데이터의 경우, 데이터를 외부 RAG에 직접 제공하는 것은 보안과 개인정보 보호 측면에서 심각한 위험을 초래할 수 있다. API 기반의 실시간 연계 방식은 민감한 데이터의 외부 노출을 최소화하며, 데이터 접근을 제어할 수 있어 보안 리스크를 줄일 수 있다.

4. 시스템 확장성과 안전성 확보: RAG가 처리해야 할 데이터의 양이 증가함에 따라 시스템의 확장성과

1) <https://scienceon.kisti.re.kr/por/intr/selectPORScienceONIntren.do?aboutCd=4>

2) 기업과 고객이 유의미한 관계를 구축하고 유지하기 위해 수행하는 일련의 모든 활동으로, 그 산출물로는 분석 보고서, 컨설팅·전략 보고서 등이 있음.

안전성 측면에서 문제가 발생할 수 있다. API 기반의 실시간 연계 방식은 필요한 데이터 크기를 유동적으로 제어함으로써 이러한 문제를 해결할 수 있다.

2.2 실시간 API 대화 기반 검색 증강 생성을 위한 추가 모듈

1. API Interface Module: 외부 정보 서비스에서 제공하는 API를 통해 데이터를 가져올 수 있도록 API 요청을 관리하고, 데이터를 수신하여 내부 RAG로 전달한다.
2. Data Integration and Preprocessing Module: API를 통해 수신된 데이터를 RAG 내부 데이터와 통합하고, 필요한 전처리 작업을 수행하는 모듈로, 외부 데이터의 품질을 보장하고, 일관된 형식으로 데이터를 제공하여 최종 결과 생성을 지원한다.
3. Caching and Query Optimization Module: 자주 요청되는 데이터를 캐싱하여 API 호출 빈도를 줄이고, API 쿼리를 동적으로 최적화하여 효율적인 데이터 검색을 지원하는 모듈로, 시스템 성능을 최적화하고 응답 시간을 단축한다.
4. Security and Privacy Module: 외부와의 데이터 통신 시 보안을 강화하고, 개인정보 보호를 위해 데이터 전송을 암호화하거나 접근을 제어하는 역할을 하여, 시스템의 신뢰성과 법적 준수를 보장한다.

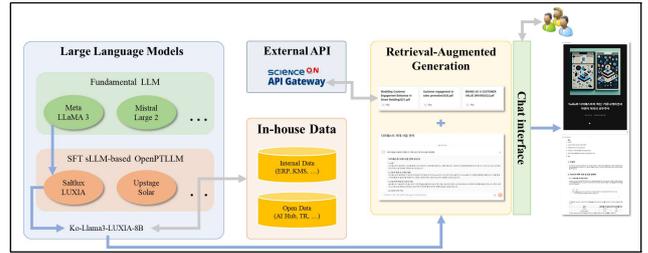
2.3 고객 인게이지먼트에의 적용 효과

1. 실시간 데이터 반영: 항상 최신 정보를 기반으로 보고서를 생성할 수 있으므로, 보고서의 신뢰성과 유효성이 높아져, 고객이 더 나은 의사 결정을 내릴 수 있다.
2. 맞춤형 보고서 생성: 고객의 특정 요구에 맞춰 적절한 외부 데이터를 API 연계 방식으로 가져와 분석함으로써, 고객이 직면한 특정 상황에 최적화된 맞춤형 전략을 제공할 수 있다.
3. 반응성과 유연성 향상: 외부 API와의 연계를 통해 고객의 긴급한 요청이나 시장의 급격한 변화에도 신속하게 대응할 수 있다.
4. 비용 효율성: 필요한 데이터만 API를 통해 실시간으로 가져와 처리하므로, 비용 효율적인 방식으로 고품질의 맞춤형 전략 보고서를 제공할 수 있다.
5. 보안 강화: 민감한 데이터를 외부에 직접 노출하지 않고 필요한 순간에만 API를 통해 가져오므로써, 보안 위험을 줄일 수 있다.
6. 확장 가능성: API를 통해 쉽게 새로운 데이터 소스를 연계할 수 있어, 지속적인 서비스 개선과 확장을 가능하게 한다.

3. 고객 인게이지먼트 분야 보고서 생성을 위한 향상된 검색 증강 생성 실험

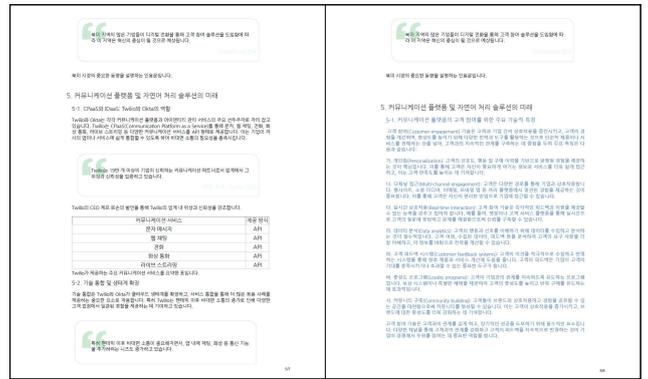
그림 1은 ㈜다이렉트를 위한 고객 인게이지먼트 분야 전략 보고서 생성을 위해 KISTI ScienceON API Gateway를 이용한 검색 증강 생성 시스템 구성도이다. 본 시스템에서 사용한 LLM은 Meta LLaMA 3를 기반으로 하여 지도 미세 조정(Supervised Fine Tuning)을 통해 개발한 솔트룩스 Ko-Llama3-Luxia-8B³⁾이며, 여기에 내부 데이터(In-house Data)를 학습시켰다. RAG를 위해 실시간으로 ScienceON API Gateway 중 산업동향보고서,

기술보고서, 특허문서 API를 통해 검색 결과를 획득하여 보고서 생성에 반영하였다.



(그림 1) 고객 인게이지먼트 전략 보고서 생성을 위한 실시간 API 대화 기반 검색 증강 생성 구성도

그림 2는 그림 1의 시스템을 통해 KISTI ScienceON 검색 결과가 보고서 일부로 자동 생성되어 추가된 예이다.



(그림 2) 고객 인게이지먼트 전략 보고서 예 (왼쪽: ScienceON API 적용 전, 오른쪽: ScienceON API 적용 후)

4. 결론

본 연구는 기존 RAG의 한계를 극복하고, 다양한 분야에서 고객의 특정 요구에 맞춘 맞춤형 정보 제공을 가능하게 하도록 실시간 API 대화 기반 RAG를 제안하였다. 이를 기반으로 고객 인게이지먼트 분야에서, 고객 맞춤형 전략 보고서를 자동 생성함으로써, 실시간 API 기반의 데이터 연계가 데이터의 최신성을 유지하고, 민감한 데이터의 보안을 강화하며, 시스템의 확장성을 유연하게 관리할 수 있다는 점을 확인하였다.

참고 문헌

[1] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. Bang, A. Madotto, and P. Fung, "Survey of hallucination in natural language generation," ACM Computing Surveys 55(12), 2023.
 [2] W. Shi, S. Min, M. Yasunaga, M. Seo, R. James, M. Lewis, L. Zettlemoyer, and W. Yih, "Replug: Retrieval-augmented black-box language models," arXiv preprint arXiv:2301.12652, 2023.
 [3] 서민택, 나승훈, 임준호, 김태형, 류희정, 장두성. "검색 증강 LLM을 통한 한국어 질의응답," 제35회 한글 및 한국어 정보처리 학술대회, 2023.

3) <https://huggingface.co/saltlux/Ko-Llama3-Luxia-8B>

5G 네트워크 자원 관리를 위한 Graph Transformer Network: 개념 연구

쉬 명하오¹, 조한규^{1,2,*}

- 1) 국립군산대학교 전자정보공학부 정보통신전파공학전공
- 2) 국립군산대학교 컴퓨터소프트웨어특성화대학 인공지능융합학과
e-mail: 962293596@qq.com, hgjo@kunsan.ac.kr

Graph Transformer Networks for Optimized 5G Network Resource Management: A Conceptual Study

Minghao Xu¹, Han-Gue Jo^{1,2,*}

- 1) School of Electronic and Information Engineering, Kunsan National University
- 2) School of Software, Artificial Intelligence Convergence, Kunsan National University

Abstract

Resource allocation in 5G networks must adapt to complex and dynamic environments. This study proposes leveraging Graph Transformer Networks (GTNs) for 5G network slicing. By modeling the 5G network as a heterogeneous graph, the approach automatically generates meta-paths to capture the intricate relationships among user equipment (UEs). Additionally, Graph Convolutional Networks (GCNs) are employed to classify nodes within the network. The proposed method aims to improve the intelligence of network resource allocation for UEs, particularly in meeting diverse network slicing requirements such as eMBB, mMTC, and uRLLC.

1. Introduction

5G networks play a vital role in supporting diverse applications through network slicing, which allows for the creation of multiple virtual networks on a shared physical infrastructure. Efficient resource management within these slices—such as spectrum and bandwidth allocation—is essential for optimizing performance. However, traditional methods often struggle to handle the dynamic and complex nature of 5G environments[1].

Graph Neural Networks (GNNs) have been utilized for network management by leveraging the underlying network topology, but their effectiveness diminishes in large-scale, dynamic settings. This study explores the application of Graph Transformer Networks (GTNs) to overcome these challenges, offering a novel approach to resource management in 5G network slices.

2. Related works

Traditional 5G resource management methods typically depend on fixed network topologies and predefined rules, making it challenging to adapt to dynamic changes in the network environment[2]. For instance, optimization algorithm-based strategies may perform well in static settings but struggle with efficiency and flexibility when faced with frequently changing traffic patterns and diverse service demands. Similarly, existing machine learning approaches often overlook the complex inter-node relationships, limiting

their ability to fully exploit the heterogeneous nature of 5G networks. These limitations hinder the ability of current solutions to meet the rapidly evolving demands and intricate resource allocation challenges inherent in 5G network slicing.

Graph Neural Networks (GNNs) have been introduced to better utilize network topology for tasks like traffic prediction and link prediction[3]. However, the application of GNNs in dynamic resource management remains constrained, particularly when dealing with the complexity of 5G slicing.

Graph Transformer Networks (GTNs) present a promising alternative by incorporating attention mechanisms that can process evolving graph structures[4]. While GTNs have demonstrated success in other domains, their potential in 5G resource management has not been thoroughly explored—an area this study seeks to address.

3. Methods

In 5G networks, different user equipment (UE) requires distinct slices of network resources (eMBB, mMTC, uRLLC) to meet specific service demands. This study aims to categorize UE nodes using GTNs to intelligently allocate network resources. To accurately represent the nodes and their interrelationships within a 5G physical infrastructure, the network is modeled as a heterogeneous graph. This graph encompasses various

types of nodes and edges, reflecting the diverse elements and connections within the 5G network.

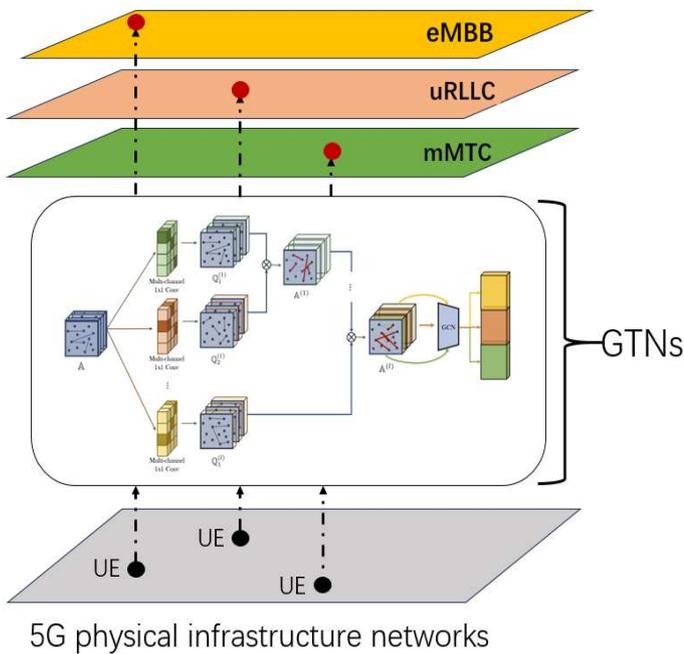


Figure 1 GTN-based 5G network resource management. The GTN figure is modified from [4].

The 5G network is modeled using three types of nodes: User Equipment (UE) representing terminal devices in the 5G network, which are the primary focus of classification; Base Station (BS) representing base stations responsible for communicating with user devices; and Edge Computing Node (EC) representing edge computing resources that provide services for low-latency applications. Within these three types of nodes, four types of edges can be established: UE to BS, BS to UE, UE to EC, EC to UE. The node features include attributes such as device type (e.g., smartphone, IoT device), transmission rate requirement, latency requirement, and connection density requirement.

To model the complex relationships within the 5G network, new graph structures called meta-paths are generated by combining different types of adjacency matrices. GTNs automatically learn and generate these meta-paths to better capture the intricate relationships between nodes. For instance, for a target UE node, GTNs might generate meta-paths like: UE → BS → UE capturing relationships between UE nodes connected through a base station; or UE → EC → UE capturing relationships between UE nodes connected via edge computing nodes.

After generating these new graph structures, GCNs are used to perform convolution operations on the graphs to extract feature representations for each node.

The inputs to the GCNs consist of the node feature matrices X and the adjacency matrices A generated by the meta-paths. Through multiple layers of convolution, the GCNs produce the final representations of each UE node, which can then be used for intelligent network resource allocation.

Finally, a classifier is employed to categorize each UE node into one of the network slices: eMBB, mMTC, or uRLLC. The classifier takes as input the node representations generated by the GCNs, and its output is the slice type label assigned to each UE node.

4. Conclusion

In this study, we propose a method utilizing GTNs for the classification of UE nodes in 5G network slicing, with the goal of optimizing network resource allocation. By modeling the 5G network as a heterogeneous graph, our approach employs GTNs to automatically generate meta-paths that capture the complex relationships among nodes. We then use GCNs to extract node features, enabling precise classification of UE nodes.

This method eliminates the need for predefined meta-paths and is capable of adaptively managing dynamic 5G network environments, thus supporting effective and intelligent network resource management. Future research could investigate the proposed concept in an experimental study with a larger-scale networks and explore its potential for other network management tasks.

Acknowledgement

This research was supported by the Korea National University Development Project.

References

- [1] Su R, Zhang D, Venkatesan R, et al. Resource allocation for network slicing in 5G telecommunication networks: A survey of principles and models[J]. IEEE Network, 2019, 33(6): 172-179.
- [2] Azimi Y, Yousefi S, Kalbkhani H, et al. Applications of machine learning in resource management for RAN-slicing in 5G and beyond networks: A survey[J]. IEEE Access, 2022, 10: 106581-106612.
- [3] Tam P, Song I, Kang S, et al. Graph neural networks for intelligent modelling in network management and orchestration: a survey on communications[J]. Electronics, 2022, 11(20): 3371
- [4] Yun S, Jeong M, Kim R, et al. Graph transformer networks[J]. Advances in neural information processing systems, 2019, 32.

실내 환경에서의 딥러닝 기반 BLE LOS/NLOS 분류

샤와나 자밀, 김소연, 변재영

조선대학교 정보통신공학과

shawana@chosun.ac.kr, kk7368@chosun.ac.kr, jypyun@chosun.ac.kr

Deep Learning based BLE LOS/NLOS Classification in Indoor Environments

Shawana Jamil, So-Yeon Kim, Jae-Young Pyun (corresponding author)
Dept. of Information and Communication Engineering, Chosun University

Abstract

This paper presents a study on the classification of Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) signals in indoor environments using Bluetooth Low Energy (BLE) Received Signal Strength (RSS). We used two deep learning algorithms LSTM and CNN-LSTM to classify unlabeled LOS/NLOS signal using BLE RSS with an accuracy of 80% and 82% respectively.

1. Introduction

One of the primary challenges in indoor positioning systems (IPS) is to select the good access point after the classification of good and bad signals for location estimation as the Received Signal Strength (RSS) values collected in indoor environments are often affected by obstacles, reflections, and multipath effects, leading to inaccuracies in location estimation. Correctly classifying these signals can greatly improve the performance of IPS. Classification of good and bad signals is termed LOS/NLOS. This paper introduces and evaluates two Deep learning Algorithms classical LSTM(Long Short-Term Memory) and CNN-LSTM (Convolutional Neural Network - Long Short-Term Memory) for distinguishing between LOS/NLOS conditions using BLE RSS.

2. Related works

Recent research has focused on various techniques such as Channel impulse response (CIR) using machine learning and deep learning [1, 2] to classify and mitigate the effects of LOS/NLOS signals. Dong et al [3] focus on using Wi-Fi RTT and RSS data for real-time identification of LOS/NLOS conditions in smartphone-based systems, employing machine learning methods to achieve high accuracy.

This paper aims to develop and validate a deep learning method for classifying unlabeled signal as LOS/NLOS using BLE RSS.

2. Methodology

2.1 Data Gathering and Preprocessing

As shown in Fig. 1, In LOS condition there is a linear relation between RSS and distance. Multiple BLE devices were deployed in a university corridor and computer lab.

Two BLE devices B4 and B9 are strategically placed Fig. 2 to capture signals in NLOS/LOS conditions respectively.

For handling missing values in acquired BLE data, we applied KNN imputation, and later various statistical features were extracted from the RSS data, such as mean, median, and standard deviation as used in [2]. These features capture essential aspects of the signal's distribution and variability, providing valuable insights for classification tasks. Additionally, the signal-to-noise ratio (SNR) was calculated using the formula (1) to quantify the quality of the signal relative to the background noise. This experiment uses *Noise floor* of -95 dbm.

$$SNR = RSS - Noise\ Floor. \quad (1)$$

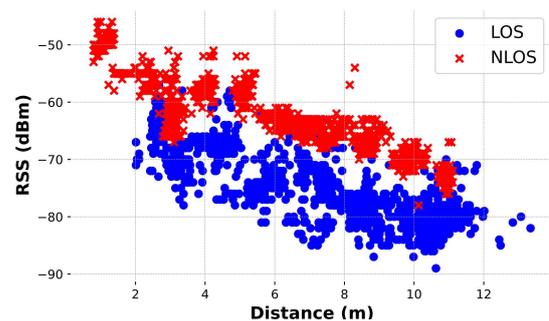


그림 1 Relationship between distance and RSS
2.2 LOS/NLOS Classification

The ability of LSTM to memorize and utilize information over long sequences makes it a good choice for distinguishing between LOS/NLOS propagation which often exhibit very small temporal variations that traditional models might overlook. Whereas the CNN-LSTM combines the strengths of CNN and LSTM networks, to process spatial and temporal data, respectively. CNN component is

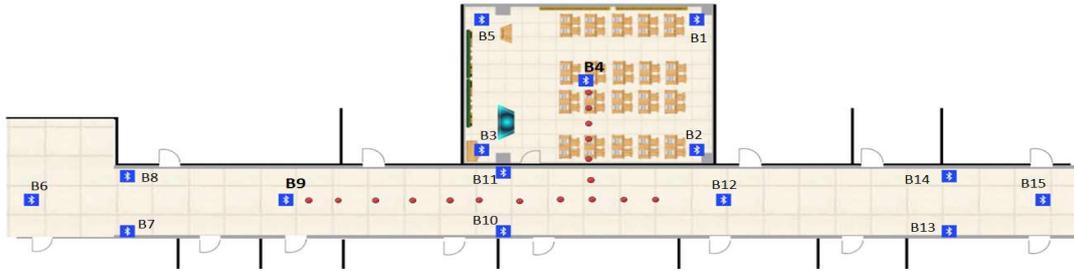


Fig. 2 Distribution of Reference points for training data collection in 10m range from B4 and B9

responsible for extracting spatial features before passing them to LSTM for capturing temporal dependencies. For experimentation, we executed 25 epochs with a batch size of 64. LSTM has 32 neurons in each of its 2 stacked LSTM layers. CNN-LSTM has 2 convolution layers with a kernel size of 3 and uses 16, and 32 filters in the first layer and second layers respectively. Both Models use the ReLU activation function and CNN-LSTM uses the softmax function in the final layer. Moreover, 2 batch normalization layers and one fully connected layer with 128 units.

3. Results

Both models show rapid initial learning and achieve final accuracy levels of around 80% and 83% respectively. In Fig. 3, the training loss of LSTM starts high at approximately 0.7 and rapidly drops to around 0.45 in the first epoch, showing significant initial improvement. However, the loss fluctuates between 0.4 and 0.2 in the

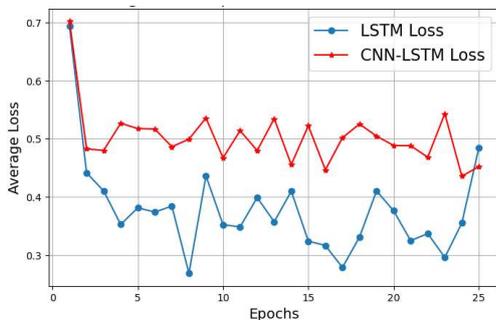


Fig. 3 Training Loss comparison of deep learning models

next epochs, indicating some instability in the training process. On the other hand, CNN-LSTM shows a slightly different behavior except for the first epoch. The training loss indicates a less sharp decline compared to the LSTM model however afterwards the loss shows less fluctuation around 0.5 with a few minor peaks, suggesting a more stable but slower learning process. The training accuracy of both models in Fig. 4 follows a similar trend, starting around 0.3 and quickly reaching 0.85, however, in the subsequent epochs, it fluctuates a lot resulting in overall accuracy of 80% and 83% respectively. The accuracy of LSTM seems slightly lower compared to the CNN-LSTM

model towards the end of training. Although LSTM model demonstrates a quicker reduction in training loss and initially faster achievement of high training accuracy, however with more fluctuation in loss indicates potential issues with overfitting or instability in learning. In contrast, the CNN-LSTM model shows a more stable training loss with less fluctuation and high final accuracy. Further evaluation of validation and test datasets needs to be done to determine their generalization capabilities.

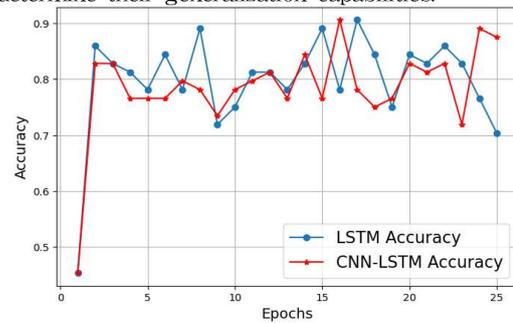


Fig. 4 Training accuracy comparison of deep learning models

Acknowledgement

This is a result of a study on the "Leaders in Industry–university Cooperation 3.0" Project supported by the Ministry of Education and National Research Foundation of Korea.

References

- [1] A. Kirmaz, D. S. Michalopoulos, I. Balan, and W. Gerstacker, "Los/nlos classification using scenario-dependent unsupervised machine learning," in 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, pp. 1134 - 1140, 2021,
- [2] T. Chang, S. Jiang, Y. Sun, A. Jia, and W. Wang, "Multi-bandwidth nlos identification based on deep learning method," in 2021 15th European Conference on Antennas & Propagation (EuCAP). IEEE, pp. 1-5, 2021.
- [3] Y. Dong, T. Arslan, and Y. Yang, "Real-time nlos/los identification for smartphone-based indoor positioning systems using wifi rtt and rss," IEEE Sensors Journal, vol. 22, no. 6, pp. 5199 - 5209, 2021.

디자인씽킹기반 SW융합 교과 설계

유소월¹, 유장웅²

전북대학교 SW중심대학사업단¹, 남부대학교 디자인공학²
e-mail : sowol@jbnu.ac.kr¹, jyu21@nambu.ac.kr²

Course Design for Software Convergence Based on Design Thinking

So Wol Yoo¹, Jang Woong You²

National Center of Excellence in SOFTWARE, Jeonbuk National University¹
Department of Design, Nambu University²

Current generations are greatly influenced by digital literacy-based sociocultural factors. However, there is a need for significant improvements in establishing a university curriculum system suitable for a software-centric society and the era of direct application of AI technologies. In response, the research team has designed a curriculum for software education utilizing design thinking, based on a literature review of domestic software education and creativity education, as well as expert group activities. As a result of the study, key elements at the learning base and teaching-learning dimensions were identified, and a standard model for related courses was designed, which is currently being applied in existing courses. Further research will be conducted to evaluate the accumulated educational outcomes in the future.

1. 서론

4차산업혁명의 시대를 살아가는 초·중·고등학생들에게 SW·AI교육이 필수적임에도 우리나라에서는 초등 17시간, 중등 34시간의 교육을 받고 있다. 고등의 경우는 정보 교과가 선택 과정으로 편성되어 있어 학생들 간의 교육 기회에 큰 차이가 발생할 수 있다. 예를 들어, 고등학교의 시도별 정보 교과의 선택 비율을 보면 대구지역에서 정보 교과의 선택 비율이 90.5%로 가장 높고, 충남지역은 17.1%로 가장 낮아 지역별로 매우 차이가 나는 것을 확인할 수 있다. 이는 모든 학생들에게 SW 교육의 기회가 공평하게 제공되지 않고 있음을 시사한다. 현재 코딩교육을 필수화하며 정보 교과 수업시간도 늘리고 있지만, 정보 교사 수도 턱없이 부족한 실정이다. 불균등한 SW교육은 단순히 현재의 문제를 넘어 상급학교로 진학하는 학생들의 SW 교육격차로 이어질 수 있는 문제점을 지적한다[1]. 따라서, 본 논문에서는 SW 교육 기회의 불균형으로 인해 대학에 진학한 학생들이 교양 수업으로서 프로그래밍 교육을 받을 때 겪는 어려움을 해소하고, 학습 효과를 극대화하기 위해 디자인씽킹을 접목한 SW융합 교과 설계를 제안한다. 이를 통해 학생들이 단순히 프로그래밍 언어의 습득을 넘어, 문제 해결 능력과 창의적 사고를 함양할 수 있는 교육적 접근 방안을 모색하고자 한다.

2. 관련 연구

2.1 국내 SW 교육 현황

국내 SW교육은 다양한 프로그램과 디바이스를 중심으로 코딩교육을 시행해 왔다. 언플러그드 코딩 교육은 사회적

기술과 논리 수학적 지능에 긍정적인 영향을 미쳤으며[2], 창의성과 미술 감상, 미술 교수 효능감이 직·간접적으로 향상된 것으로 보고되었다[3].

반면, 최근 연구자들은 공감력과 협동 능력에는 보완이 필요하고, 순차적 문제해결 이상의 창의적 문제해결을 위한 교수학습 수단의 필요성에 주목하고 있다.

2.2 디자인씽킹을 활용한 SW교육의 효과

디자인씽킹을 소프트웨어 교육에 접목했을 때, 다양한 긍정적인 효과를 확인할 수 있다. 디자인씽킹 프로세스와 머신러닝 학습 프로그램을 결합하여 창의적 문제 해결 능력을 향상시키는 데 성공하였으며, 지식, 사고, 기능 및 기술 이해, 확산적 사고, 논리적 및 비판적 사고, 동기적 요소 등 여러 측면에서 유의미한 향상이 나타났다[4]. 또한, 창의적 문제 해결력과 창의적 성향을 개발하기 위한 프로세스 기반 교육과 프로그램 지원이 교수자를 포함한 학습자의 인내, 호기심, 자신감 등 하위 문제 영역에서도 긍정적인 성과를 보였다[5].

2.3 학습 효과 가설

디자인씽킹을 SW교육과정에 연계하여 실습이 병행되는 창의적 사고와 문제해결(Creative Thinking and Problem Solving), 융합적 사고와 창의성(Convergent Thinking and Creativity), 디자인사고(Design Thinking), 미디어와 창의성(Media and Creativity), 창의성과 코딩(Creativity and Coding) 5개 교과들의 공통요소와 창의적 글쓰기(Creative Writing), 창의적 리더십(Creative Leadership),

예술과 창의성(Art and Creativity) 창의력 향상 3개 교과의 학습효익에 대한 전문가집단 쇼카드 분석 결과로 표 1과 같이 4개의 가설이 추출되었다.

(표 1) 학습 효익에 대한 가설

구분	내용
가설 1	학습자의 창의력 증진에 효과가 있다.
가설 2	알고리즘 학습에 효능이 있다.
가설 3	학습자의 학습효과를 높일 수 있다.
가설 4	PBL 학습을 통해 협업 능력이 고양될 수 있다.

2.4 학습효능 지속을 위한 요소

상기 가설들을 확인하기 위해서는 학습기반과 교수학습 영역에서 각각 최소 5개 요소가 요구된다. 사전연구에서 진행된 12인의 관련교과 교양교원의 마인드맵 구조화 결과 도출된 10대 요소들과 N대학 관련 요소별 확보 현황은 표 2와 같이 진단 및 예측 분석, 맞춤형 학습경험 제공, 후속지원 및 성과 확산 부문에 보완이 필요하다.

(표 2) N대학의 학습효능 지속성 확보 요소별 확보 현황

학습기반	현황	교수학습	현황
교수학습관리지원 플랫폼	●	교수학습 지원	●
단계별 교양교육과정체계	●	학습경험 Mining 효익	◎
시스템 및 디바이스	◎	학습효능 향상	●
진단 및 예측 분석	○	맞춤형 학습 경험 제공	○
성과 분석 환류 체계	◎	후속지원 및 성과확산	○

2.5 교과설계

N대학교의 SW기초 창의적 사고와 코딩 교과는 소프트웨어 교양필수 교과목으로 모든 학생들이 이수하고 있다. 따라서 SW수준과 상관없이 해당 교양과목의 수강자들이 적극적으로 참여할 수 있는 동기를 제공하는 교수 설계가 필요하다고 판단하였다.

이에 컴퓨팅사고와 디자인씽킹을 적용한 창의적 문제해결 방법을 학습하고, 팀프로젝트에 적용한 코딩실습을 통해 최종 결과물을 얻도록 설계하였다. 2주차부터 3주차까지는 창의성의 개념 및 접근법에 대해 학습하며, 팀 단위로 일상생활에서 발생할 수 있는 문제점들을 발견하는 데이터 수집단계를 진행한다. 4주차부터 7주차까지는 수집한 데이터를 기반으로 문제점을 정의하고 아이디어를 도출할 수 있는 문제해결능력 향상 단계를 진행한다. 9주차부터 11주차까지는 문제점을 해결할 수 있는 알고리즘을 설계하고 구현하는 탐구와 실험 과정을 진행한다. 프로토타입을 제작해봄으로써 실제로 적용가능한지 확인하고, 수정할 부분은 다시 진행한다. 12주차부터 13주차까지는 해당 아이디어를 적용한 코딩을 진행하며, 14주차에는 팀 아이디어를 발표하고 평가하도록 설계하였다.

(표 3) 디자인씽킹기반 SW융합 교과 설계

주차	주제	컴퓨팅사고	디자인씽킹	팀프로젝트	코딩
1	이해와 도입			팀 구성 역할 분담	블록코딩 소개
2	창의성 개념 및 접근법	데이터수집	문제발견	계획 수립	코딩 환경 기동과 사용
3					기초 코딩 프로젝트
4	문제해결능력	분해	문제정의	문제 인식	반복 조건문

5				아이디어 발전을 위한 토론	재어블록 게임형 프로젝트
6	향상	추상화 & 패턴인식	아이디어	문제 해결 알고리즘 사고 학습	인터랙티브 스토리텔링 프로젝트
7				문제 해결 방법 논의 및 실습	
8	중간고사				
9	탐구와 실험	알고리즘 설계	프로토타입 제작	프로젝트 구현 계획 수립	반수 개수와 리스트 활용
10	합업과 탐구	알고리즘 구현		프로젝트 구현	
11					사물라이선 프로젝트 제작
12	비판적 사고와 혁신	실험 및 테스트	실험 및 테스트	프로젝트 테스트 후 피드백 및 개선 적용	
13					
14	학습 경험 공유			프로젝트 발표 및 평가	
15	기말고사				

3. 결론 및 향후 연구

전통적으로 직무능력과 사회적 활동능력 향상을 위해 대학교양에서는 프로세스기반 문제해결과 소통형 학습, PBL 교육을 운영해왔다. 또한 소프트웨어 중심사회로 진입하면서 가속화된 코딩교육과 창의중심 사고혁신은 아이디어 선과정 이상의 디자인씽킹 교육방법론을 수용하고 있다. 최근 AI기술의 발전은 생활과 직무에서 보조적 역할을 벗어나 실제적 활용 빈도가 높아졌고 급속한 질적 성장속도로 인해 현 대학생들의 근미래 사회직무환경은 프롬프트 기반의 알고리즘 활용에 개인과 집단의 창의력을 어느 정도 반영하는가에 중요도가 실릴 것이라 예측되고 있다. 본 연구에서는 단인 교과과정 개발의 한계점은 있으나 가장 융합적 교육교과로 활용되는 창의적 코딩 교과유형의 교과설계를 통해 학습자의 사용환경과 근본적 학습목표 달성을 위한 교과 기준을 제시코자 한다.

참고 문헌

- [1] 최문석 외, 디지털인재 양성을 위한 초등학교의 보편적 정보교육 도입의 필요성 연구, 정보교육학회논문지, 26[3], pp.209~218, 2022
- [2] 신원애 외, 누리과정 연계 언플러그드코딩프로그램 개발 및 효과, 한국유아교육연구, 26[1], pp.75~100, 2024
- [3] 박은주, 탐색, 표현, 감상의 통합적 아동미술 수업이 예비유아교사의 창의성과 미술감상능력 및 미술교수효능감에 미치는 효과 및 변화과정 탐색, 학습자중심교과교육학회지, 24[1], pp.413~432, 2024
- [4] 김영동 외, 창의적 문제해결력 향상을 위한 디자인 씽킹 프로세스 기반 머신러닝 학습프로그램 개발과 적용, 초등교육연구논총, 40[1], pp.161~180, 2024
- [5] 박재현 외, 통합적 메이커교육이 예비교사의 창의적 문제해결력 및 창의적 성향에 미치는 영향, 통합교육과정연구, 18[3], pp.157~180, 2024

대형 언어 모델을 위한 효율적인 연합 학습 방법에 대한 연구

시바니 산제이 콜레카르, 김경백
 전남대학교 인공지능학과
 shivanikolekar@gmail.com, kyungbaekkim@jnu.ac.kr

A Study of Efficient Federated Learning for Large Language Models

Shivani Sanjay Kolekar, Kyungbaek Kim
 Dept. of Artificial Intelligence Convergence,
 Chonnam National University Gwangju, South Korea

Abstract

The main goal of this paper is to explore energy-efficient approaches to federated learning (FL) for large language models (LLMs). As LLMs continue to grow in size and complexity, traditional centralized training poses significant challenges in terms of computational costs, data privacy, and energy consumption. Federated learning offers a decentralized solution that addresses these issues by distributing training across multiple devices while preserving data privacy. This paper reviews the key components of FL systems, including benchmarks, back-propagation-free optimization methods, and fine-tuning techniques, all aimed at reducing energy consumption and improving efficiency. We discuss the challenges of data heterogeneity, privacy, and computational resource management in FL-LLM systems and provide insights into future directions for making federated LLM training more energy-efficient and scalable.

1. Introduction

Large language models (LLMs) have demonstrated remarkable capabilities in processing and generating human-like text[14], making them indispensable in diverse domains, from education[3] to healthcare[1]. However, to maximize the utility of these models in specific fields, fine-tuning is essential. Fine-tuning adapts pre-trained models with domain-specific data, enabling them to perform better on specialized tasks by incorporating domain knowledge. This process enhances the reliability of outputs, improves fairness, and mitigates issues like hallucinations, where models generate inaccurate or misleading information.

Despite these benefits, fine-tuning LLMs presents significant challenges. The computational resources required for fine-tuning models with billions of parameters are substantial, often necessitating advanced infrastructure. Additionally, access to domain-specific data, which is crucial for fine-tuning, can be limited due to privacy concerns, particularly in sensitive sectors like healthcare and finance. These issues highlight the need for efficient and privacy-preserving fine-tuning methods. Federated learning (FL) offers a promising solution by enabling multiple entities to collaborate in fine-tuning models without sharing raw data, ensuring privacy while benefiting from collective learning. Techniques such as model decomposition and back-propagation-free strategies further enhance the efficiency of federated fine-tuning by optimizing key components of the model and reducing computational costs. Model decomposition allows certain parts of the model to remain static while others are updated, minimizing resource use. Back-propagation-free approaches provide alternative optimization techniques that avoid gradient-based computations, significantly lowering the overall cost.

To evaluate the effectiveness of these approaches, new benchmarks have been developed. These benchmarks are critical for assessing the real-world applicability of federated

LLM fine-tuning across distributed and heterogeneous data environments.

In this paper, we discuss current methodologies for federated LLM fine-tuning, including model decomposition and back-propagation-free approaches, and examine existing benchmarks that assess their performance. We also explore future directions for improving scalability, efficiency, and privacy to guide the development of more effective federated fine-tuning frameworks for LLMs in practical applications.

2. Resource-Efficient FL-LLM Systems

We summarize the FL-LLM models and benchmarks related to this review concisely in the Table 1.

i. PEFT Strategies

Parameter-Efficient Fine-Tuning (PEFT) is a strategy designed to tailor LLMs for specific downstream tasks. FedAdapter [5] proposes progressively modifying the adapter configuration to efficiently identify the most effective setup. FeS [6] integrates bias-only prompt learning with pseudo-labeling in a structured curriculum, achieving substantial accuracy with minimal data labeling. FP-FL [7] introduces a soft-label enhanced federated tuning, incorporating LoRA tuning to diminish both computational and communication expenses. FedPrompt [8] explores the additional advantages of prompt learning, particularly its potential to allow for larger differential privacy budgets. Malaviya et al. [9] analyzed the effectiveness of various PEFT

Resource Efficient LLM Systems following Federated learning mechanism	
Framework & Benchmark	Flower [10], FedML [11], FATE [12], Federatedscope-llm [13], Woisetschlger et al. [15]
PEFT-based Approaches	FedAdapter [5], FeS [6], FP-FL [7], FedPrompt [8], Malaviya et al. [9]
Model Decomposition	FedOBD [17], FedBFPT [16]
Backprop-free Approaches	FwdLLM [18] FedKSeed [19]

et al. [9] analyzed the effectiveness of various PEFT

methods under different non-IID scenarios and varying client fractions.

ii. Frameworks and Benchmarks

Significant progress has been made in federated learning frameworks, including Flower [10], which supports federated fine-tuning of LLMs, FedML [11] with its FedLLM integration, FATE [12] enhancing training efficiency with parameter-efficient tuning, FederatedScope-LLM [13] offering a comprehensive package for federated fine-tuning, and Woissetschlager et al. [15], who evaluated the potential of edge computing for large federated models.

iii. Model Decomposition

Decomposing a large FM into several sub-models is a straightforward yet effective approach towards practical FL. FedOBD [17] takes an innovative approach by decomposing large scale models into semantic blocks, enabling FL participants to selectively upload quantized blocks to the FL server for aggregation. FedBFPT [16] adopts a strategy where only a portion of BERT's layers are trained on the client side, with the number of layers involved gradually increasing.

iv. Backpropagation-free Approaches

Back-propagation (BP) is challenging in FL due to high computational and memory costs on edge devices. Zeroth-order optimization offers an alternative by optimizing large FMs without relying on BP gradients. FwdLLM [18] integrates zeroth-order optimization in FL using forward gradients, but reduces gradient estimations by updating only intrinsic dimensions. It enables training large LLMs on mobile devices. FedKSeed [19] avoids forward gradient transmission by using a gradient accumulator to compute local gradients.

3. Conclusion and Future Directions

In this paper, we reviewed energy-efficient federated learning (FL) for large language models (LLMs), focusing on reducing computational costs, data privacy risks, and energy consumption. While FL shows promise, challenges like communication overhead, data heterogeneity, and resource management remain. Future research should prioritize improving communication efficiency, optimizing for diverse data environments, and enhancing privacy preserving techniques, along with hardware-aware optimizations and robust benchmarks to ensure scalable, energy-efficient FL-LLM systems.

Acknowledgement

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government (MSIT). This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2024-00437718) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation)

References

[1] Google, "A large language model from Google research, designed for the medical domain," 2023.

- [2] Yu, Tao, Eugene Bagdasaryan, and Vitaly Shmatikov. "Salvaging federated learning by local adaptation." arXiv preprint arXiv:2002.04758 (2020).
- [3] Latif, Ehsan, et al. "Knowledge distillation of llm for education." arXiv preprint arXiv:2312.15842 (2023).
- [4] Alex de Vries. The growing energy footprint of artificial intelligence. *Joule*, 7(10):2191–2194, 2023.
- [5] Cai, Dongqi, et al. "Efficient federated learning for modern nlp." Proceedings of the 29th Annual International Conference on Mobile Computing and Networking, 2023.
- [6] Cai, Dongqi, et al. "Federated few-shot learning for mobile nlp." Proceedings of the 29th Annual International Conference on Mobile Computing and Networking, 2023.
- [7] Jingang Jiang, Xiangyang Liu, and Chenyou Fan. Low-parameter federated learning with large language models. arXiv preprint arXiv:2307.13896, 2023.
- [8] Zhao, Haodong, et al. "Fedprompt: Communication-efficient and privacy-preserving prompt tuning in federated learning." ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023.
- [9] Shubham Malaviya, Manish Shukla, and Sachin Lodha. Reducing communication overhead in federated learning for pre-trained language models using parameter-efficient finetuning. In Conference on Lifelong Learning Agents, pages 456–469. PMLR, 2023.
- [10] Beutel, Daniel J., et al. "Flower: A friendly federated learning research framework." arXiv preprint arXiv:2007.14390 (2020).
- [11] He, Chaoyang, et al. "Fedml: A research library and benchmark for federated machine learning." arXiv preprint arXiv:2007.13518 (2020).
- [12] Fan, Tao, et al. "Fate-llm: A industrial grade federated learning framework for large language models." arXiv preprint arXiv:2310.10049 (2023).
- [13] Kuang, Weirui, et al. "Federatedscope-llm: A comprehensive package for fine-tuning large language models in federated learning." Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024.
- [14] Winata, Genta Indra, et al. "Language models are few-shot multilingual learners." arXiv preprint arXiv:2109.07684 (2021).
- [15] Herbert Woissetschlager, Alexander Isenko, Shiqiang Wang, Ruben Mayer, and Hans-Arno Jacobsen. Federated fine-tuning of llms on the very edge: The good, the bad, the ugly. arXiv preprint arXiv:2310.03150, 2023.
- [16] Xin'ao Wang, Huan Li, Ke Chen, and Lidan Shou. "FEDBFPT: An efficient federated learning framework for BERT further pre-training." Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23, 2023.
- [17] Chen, Yuanyuan, et al. "FedOBD: Opportunistic block dropout for efficiently training large-scale neural networks through federated learning." arXiv preprint arXiv:2208.05174 (2022).
- [18] Xu, Mengwei, et al. "FwdLLM: Efficient FedLLM using Forward Gradient." arXiv preprint arXiv:2308.13894 (2023).
- [19] Zhen Qin, Daoyuan Chen, Bingchen Qian, Bolin Ding, Yaliang Li, and Shuiguang Deng. Federated full parameter tuning of billion-sized language models with communication cost under 18 kilobytes. arXiv preprint arXiv:2312.06353, 2023.

ASD 조기 예측을 위한 연합학습 기반 시선 추적 이미지 분류 모델

신현호, 이재민, 전홍우
 한국과학기술정보연구원 미래기술분석센터
 e-mail : {shinhh9554, jmlee, hw.chun}@kisti.re.kr

Federated Learning based Eye-tracking Image Classification Model for Early Detection of ASD

Hyeonho Shin, Jae-Min Lee, Hong-Woo Chun
 Korea Institute of Science and Technology Information,
 Future Technology Analysis Center

요약

자폐 스펙트럼 장애(ASD)는 반복적인 행동과 사회적 의사소통 문제를 특징으로 하는 신경 발달 장애로, 조기 발견과 치료가 예후에 큰 영향을 미친다. 최근에 시선 이동 패턴을 분석하여 ASD를 조기에 발견하려는 연구가 활발히 진행되고 있다. 그러나 의료 데이터로서 환자의 시선 추적 이미지는 수집과 공유에 제약이 있다. 이를 극복하기 위해, 본 연구에서는 ASD 환자의 시선 추적 이미지를 활용한 연합학습(Federated Learning) 기반 이미지 분류 모델을 제안하였다. 실험 결과, 연합학습 모델 성능은 정확도 0.85, F1 스코어 0.84를 보였다. 이를 통해 ASD 조기 예측에서 연합학습이 효과적인 접근법을 입증하였으며, 의료 데이터의 개인정보 보호와 분산 학습의 가능성을 동시에 제시하였다.

1. 서론

자폐 스펙트럼 장애(Autism Spectrum Disorder, ASD)는 신경 발달 장애의 일종으로, 사회적 상호작용과 의사소통에 장애를 겪고, 반복적인 행동을 보이는 것이 특징이다[1-2]. ASD의 조기 발견과 치료는 매우 중요한데, 만 5세 이전에 발견하고 치료를 시작하면 예후가 크게 달라지기 때문이다[3]. 최근 딥러닝을 활용하여 ASD 환자의 시선 이동 패턴을 분석하고, 이를 통해 ASD를 조기에 발견하고 병증의 정도를 파악하려는 연구들이 활발히 진행되고 있다. 이는 시선의 이상 패턴이 일반적으로 ASD 환자의 주요 특징으로 인식되기 때문이다[1].

하지만 환자를 대상으로 한 시선 추적 이미지 데이터는 의료 데이터로 수집과 공유에 제약이 있다[2-3]. 본 연구에서는 이러한 한계를 보완하기 위해 ASD 환자의 시선 추적 이미지 데이터를 이용해 연합학습(Federated Learning)을 적용한 이미지 분류 모델을 제안한다. 본 연구는 이를 통해 연합학습의 효용성을 검증하고, 분산된 데이터 환경에서도 ASD 예측 모델 학습이 가능함을 입증하고자 한다.

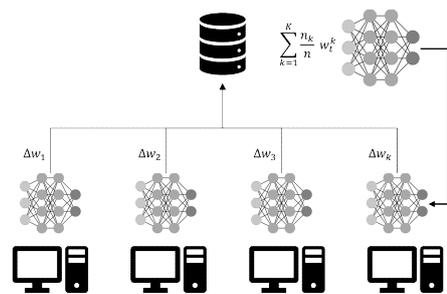
2. 관련 연구

최근 ASD 연구에 사용되는 데이터의 개인정보 보호를 위해 연합학습이 활용되고 있다. 이에 [2]와 [3]에서는 연합학습을 활용하여 개인정보를 보호하면서도 높은 성능을 달성하는 ASD 예측 모델을 구축하였다. 그러나 시선 패턴 분석을 통한 ASD 예측 연구는 아직 부족한 상황이다. 따라서 본 연구에서는 연합학습을 적용하여 시선 패턴을 분석하고, 이를 통해 ASD를 예측하는 모델을 구축하고자 한다.

3. 제안 방법

3.1 연합학습 환경 구성

연합학습은 데이터를 직접 공유하지 않고, 각 클라이언트에서 로컬 학습한 가중치를 서버로 전송해 글로벌 모델을 학습하는 방식이다. 글로벌 모델은 클라이언트의 학습 결과를 통합하여 업데이트되며, 이 과정을 Round라고 한다[2]. 연합학습 프레임워크는 그림 1과 같다.



(그림 1) 연합학습 프레임워크

각 클라이언트에서 사용되는 신경망은 CNN 구조로, 합성곱층, 최대 풀링, 드롭아웃을 통해 이미지를 처리하고 최종 클래스를 예측한다. 연합학습 환경에서 서버는 TITAN RTX GPU 1대를 장착한 워크스테이션으로 구성되며, 클라이언트는 총 2대로 구성되어 있다. 첫 번째 클라이언트는 RTX 6000 ADA GPU 2대를, 두 번째 클라이언트는 RTX A6000 GPU 1대를 장착하고 있다. 모든 클라이언트는 각 라운드에 참여하며, 각 클라이언트가 보유한 로컬 데이터 샘플의 개수는 동일하다.

3.2 로컬 및 전역 모델 훈련 과정

연합학습에서 각 로컬 클라이언트는 ASD 대상 시선 추적 이

미지를 활용하여 CNN 기반의 이미지 분류 모델을 학습하고, 학습이 완료되면 모델의 파라미터를 서버로 전송한다. 서버는 전송 받은 파라미터들을 수식 1과 같이 평균화하여 글로벌 모델을 업데이트한다. 글로벌 모델 업데이트는 각 클라이언트의 데이터 Sample 수에 비례하여 가중 평균을 계산하는 방식으로 이루어지며, 라운드가 반복될수록 모델의 파라미터 값이 수렴하고, 글로벌 모델의 정확도 또한 계속 증가하여 최종 정확도에 수렴하게 된다.

$$w_{i+1} \leftarrow \sum_{k=1}^K \left(\frac{n_k}{n} w_i^k \right)$$

(수식 1) 글로벌 모델 업데이트 수식

4. 실험

4.1 실험 환경

본 연구에서는 시선 추적 이미지를 통해 ASD와 TD(전형적 발달) 아동을 분류하기 위하여 [1]에서 구축한 547개의 시선 추적 이미지 데이터셋을 활용하였다. ASD 클래스에는 219개, TD 클래스에는 328개의 이미지가 있으며, ASD 진단은 CARS(Childhood Autism Rating Scale) 점수를 기반으로 평가되었다. 본 연구에서는 ASD 유무뿐만 아니라 그 정도를 파악하기 위해 CARS 점수를 기준으로 ASD 클래스를 경증(CARS < 30), 중등도(30 ≤ CARS < 36), 중증(CARS ≥ 36)의 세 가지 클래스로 구분하였다. 구분된 데이터 수는 표 1과 같다. [1]에서 구축한 데이터 수는 CNN 모델을 학습하기에는 충분하지 않아 데이터 증강을 적용하였다. 각 클라이언트에 배정된 학습 데이터 Sample 수는 2,704개이다.

(표 1) 시선 추적 이미지 각 Class 별 데이터 수

Class	Train	Test
ASD 경증	88	10
ASD 중등도	78	9
ASD 중증	30	4
TD	295	33

모델 성능 비교 실험은 중앙 집중식 학습과 연합학습에 대해 모두 진행하였다. 두 학습 방법 모두 Epoch는 20회, Batch 크기는 16으로 설정하였으며, 연합학습은 20회 Round를 수행하였다. 손실함수로는 교차 엔트로피 함수를 사용하였고, Optimizer는 모두 Adam을 사용하였다.

4.2 실험 결과

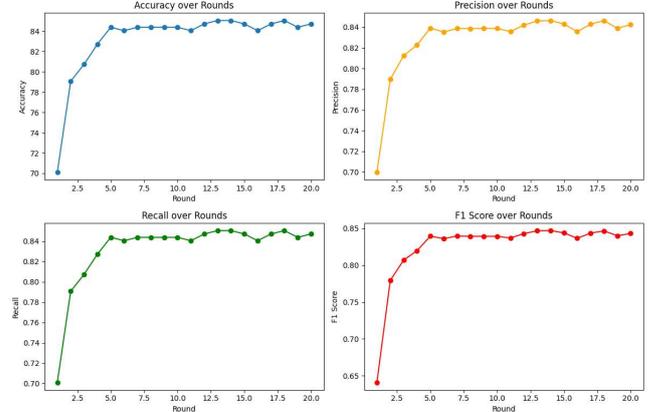
실제 연합학습을 적용한 분류 모델의 성능을 평가하기 위해 전체 데이터를 모두 학습한 중앙 집중형 모델과 연합학습으로 학습한 모델의 성능을 평가하였다.

(표 2) 학습 방법에 따른 모델 성능

학습 방법	ACC	F1	Precision	Recall
중앙집중	0.92	0.89	0.94	0.85
연합학습	0.85	0.84	0.84	0.85

중앙집중형 모델의 성능은 표 2와 같이 정확도 0.92, F1 스코어 0.89를 보였다. 연합학습 방법을 적용한 글로벌 모델의 성능은 그림 2와 같이 약 5라운드에서 기존 방법과 유사한 성능을 보이는 것으로 확인하였다. 최종 학습된 모델의 성능은 표 2와 같이 정확

도는 0.85, F1 스코어는 0.84의 성능을 보였다. 연합학습이 기존 학습 방법 대비 약 95% 이상의 성능을 달성함을 확인할 수 있었다. 위 실험 결과는 연합학습이 중앙집중형 모델 대비 약간의 성능 저하가 있지만 분산된 데이터 환경에서 개인정보 보호를 유지하며 일정 수준 이상의 성능을 도출할 수 있음을 보여준다.



(그림 2) 연합학습 Round 별 모델 성능

5. 결론

본 연구에서는 ASD와 TD 아동 분류를 위해 시선 추적 이미지에 기반한 연합학습 모델을 제안하였다. 모델은 의료 데이터의 개인정보 보호를 보장하면서도 중앙 집중식 학습과 유사한 성능을 달성했다. 비록 연합학습 모델이 기존 학습 방법 대비 약간의 성능 저하를 보였지만, 분산 데이터 환경에서도 학습이 가능함을 입증하여 그 유용성을 확인했다. 향후 연구에는 모델 성능 향상을 위해 여러 네트워크 및 학습 기법을 적용하고, 실제 임상 환경에서의 적용 연구를 진행할 예정이다.

참고 문헌

[1] Carette, R., Elbattah, M., Cilia, F., Dequen, G., Guerin, J. L., Bosche, J., "Learning to Predict Autism Spectrum Disorder based on the Visual Patterns of Eye-tracking Scanpaths", In Proceedings of the 12th International Conference on Health Informatics, 2019, pp. 103-112.

[2] H. Shamseddine, S. Otoum and A. Mourad, "On the Feasibility of Federated Learning for Neurodevelopmental Disorders: ASD Detection Use-Case", GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 1121-1127.

[3] Farooq, M. S., R. Tehseen, M. Sabir., "Detection of autism spectrum disorder (ASD) in children and adults using machine learning", Scientific Reports, Vol.13, No.9605, June 2023. DOI: 10.1038/s41598-023-35910-1.

감사의 글

본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건 의료기술연구개발사업 지원에 의하여 이루어진 것임(과제고유번호: HR22C1605)

다중 양식 융합과 개선된 E-MTN을 활용한 비디오 폭력 탐지 성능 향상

나광호, 최경주

충북대학교 소프트웨어학과

e-mail : gh369ho@naver.com, kjcheoi@cbnu.ac.kr

Enhancing Video Violence Detection Performance Using Multi-modal Fusion and E-MTN

Gwangho Na, Kyung Joo Cheoi

Dept. of Computer Science, Chungbuk, National University

요약

본 논문에서는 다중 양식 융합과 개선된 E-MTN을 활용하여 비디오 폭력 탐지 성능을 향상한 새로운 모델을 제안하였다. 제안된 모델은 기존의 이중 양식 모델인 TEVAD를 확장한 것으로, 시각, 음성, 문자 정보의 3중 양식을 활용하였고, 다양한 실험을 통해 모든 양식의 특징을 통합한 후에 비디오 스니펫 간의 다중 시간 스케일 특징 추출 네트워크(MTN)를 적용하는 것이 가장 우수한 성능을 보임을 확인하였으며, 특히 각 특징을 단순히 연결하는 방법이 가장 효과적임을 입증하였다. 또한, 개선된 MTN을 사용하여 성능이 향상됨을 확인하였다.

1. 서론

최근 몇 년간 비디오 콘텐츠의 소비가 급증하면서 폭력적인 장면을 포함한 비디오를 자동으로 감지하는 기술에 대한 필요성이 증가하고 있다. 기존의 폭력 탐지 연구는 주로 시각 정보에 의존하여 개발되었다. 하지만 이러한 접근 방식은 복잡한 실제 상황에서 발생하는 다양한 형태의 폭력을 정확하게 탐지하는 데 한계가 있다. 비디오를 정확하게 해석하기 위해서는 시각 정보뿐만 아니라 다중 양식(Multi-modal)을 고려할 필요가 있다. 시각 정보만으로는 부족한 의미론적 의미를 포함하고 있으므로 다중 양식을 비디오를 효과적으로 해석하는 데 도움을 줄 수 있다.

최근의 비디오 폭력 탐지 연구들은 완전히 라벨이 달린 데이터가 아닌, 부분적으로 또는 간접적으로 라벨이 제공된 데이터를 사용하여 모델을 학습하는 접근법인 약한 지도 학습(weakly supervised learning)[1]을 사용한다. 약한 지도 학습 방법은 대부분 다중 인스턴스 학습(Multiple Instance Learning, MIL) 프레임워크를 기반으로 하는데, MIL은 클립 내의 각 구간에 대한 세부 라벨은 사용하지 않고, 비디오 클립 전체에 대한 라벨을 사용해 클립 내의 프레임 혹은 짧은 구간마다 라벨을 추정한다.

Chen 등[2]은 기존의 시각 특징만으로는 파악하기 어려운 의미론적 정보를 포착하기 위해 생성한 문자 특징을 시각 특징과 결합한 문자 기반 비디오 이상 탐지 시스템인 TEVAD(Text Empowered Video Anomaly Detection)를 제안하였다. TEVAD는 비디오 내의 의미론적 의미를 포착한 문자 정보를 함께 사용함으로써 기존의 시각적 특징 기반 방법보다 향상된 성능을 보여주었다.

본 논문에서는 시각 정보와 문자 정보를 활용한 TEVAD를 확장하여 시각 특징만으로는 추출하기 어려웠던 의미론적 의미를 비디오 캡처링 모델을 사용해 생성한

문자(캡션)를 통해 포착하고, 음성 정보를 융합하는 통합적인 접근 방법을 제안한다. 기존 TEVAD와는 음성 정보를 추가로 사용한다는 점과 특징 융합의 방법 및 시기가 다르다는 차이가 있다. 또한 다중 시간 스케일에서 시간 특징을 더욱 효과적으로 추출하기 위해 기존의 MTN(Multi-scale Temporal Network)을 수정한 개선된 E-MTN을 제안한다. 이러한 차이로 비디오 내 폭력을 탐지하는 성능을 극대화시켰다.

2. 제안 방법

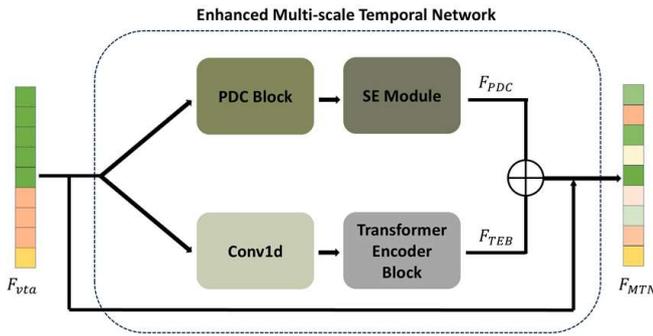
2.1 제안된 구조

제안된 모델의 전체 구조를 설명하면 다음과 같다. 먼저 비디오가 모델에 입력되면 T개의 스니펫(snippets)으로 분할되고, 분할된 각 스니펫에 대해 시각 특징 F_v , 음성 특징 F_a , 문자 특징 F_t 를 추출한다. 이렇게 추출된 각 특징들은 연결(Concatenation) 방법을 통해 하나의 특징으로 융합되어 F_{vta} 를 생성한다. 이후 융합된 특징은 개선된 E-MTN(Enhanced Multi-scale Temporal Networks)을 거쳐 다중 시간 스케일 특징 F_{E-MTN} 을 생성한다. 이렇게 생성된 F_{E-MTN} 은 각 스니펫의 특징 크기(magnitude) 계산에 사용되어진다. 정상과 폭력 비디오로부터 만들어진 상위 k개의 가장 큰 특징 크기가 스니펫 수준에서 폭력인지 아닌지를 탐지하기 위한 폭력 탐지 분류기의 훈련을 위해 전달되어 폭력 탐지를 위한 2진 분류기를 훈련한다.

2.2 개선된 E-MTN

기존의 MTN[2]은 다양한 시간 스케일에서 시간 특징을 추출하기 위해 피라미드 확장 컨볼루션(PDC) 블록과

Non-local 블록(NLB)을 사용하였다. 본 논문에서는 SE 모듈(Squeeze and Excitation Module)을 추가하고, NLB를 트랜스포머 인코더 블록(TEB)으로 대체한 개선된 E-MTN을 제안한다, 그림 1은 개선된 E-MTN의 전체적인 구조도를 보여준다. F_{E-MTN} 을 추출하는 절차는 다음과 같다. 먼저 융합된 특징 F_{vta} 는 PDC 블록과 SE 모듈을 통과하여 F_{PDC} 를 생성한다. 동시에 F_{vta} 는 1D-컨볼루션을 거친 후 TEB를 통과하여 F_{TEB} 를 생성한다. 이후 F_{PDC} 와 F_{TEB} 를 연결(Concatenation)하고, 원래 특징인 F_{vta} 를 더하여(Addition) 최종 출력 F_{E-MTN} 을 생성한다.



(그림 1) 개선된 E-MTN 구조도

3. 실험 및 결과

제안하는 시스템의 성능을 평가하기 위해 다중 양식을 제공하는 XD-Violence[3] 데이터셋을 사용하였다. 표 1은 XD-Violence 데이터셋에 대한 프레임 수준의 AP 성능을 비교한 결과이다. 본 연구에서 제안한 다중 양식 접근법의 효과를 평가하기 위해, 성능이 우수한 기존의 단일 양식 및 이중 양식 모델들과 성능을 비교하였다. 본 연구에서 제안된 다중 양식 접근법은 시각, 음성, 문자 정보를 모두 결합하여 사용함으로써 83.9%의 AP 성능을 기록하였다.

(표 1) XD-Violence 데이터셋에 대한 프레임 수준 AP 성능 비교

Method	Modality	AP(%)
RTFM (2021) [4]	V	77.8
TEVAD (2023) [2]	V, T	79.8
Zhang et al. (2023) [5]	V, A	81.4
UR-DMU (2023) [6]	V, A	81.7
Ours (Original MTN)	V, T, A	82.2
Ours (Enhanced MTN)	V, T, A	83.9

4. 결론

본 논문에서는 TEVAD를 확장하여 시각, 음성, 문자 정보를 모두 활용한 다중 양식 비디오 폭력 탐지 시스템을 제안하였다. 다양한 실험을 통해, 모든 양식의 특징을 하나의 통합된 특징으로 융합한 후 MTN을 거치도록 하는 방법이 가장 우수한 성능을 보임을 확인하였다. 다양한 특징 융합 방법 중에서는 각 특징들을 단순히 이어 붙이

는 연결(Concatenation) 방법이 가장 효과적임을 입증하였으며 다중 양식을 사용하는 경우가 기존 단일 양식 또는 이중 양식을 사용하는 경우보다 성능이 더 우수함을 확인하였다. 또한 대부분의 실험에서 기존 MTN을 사용한 경우보다 제안된 개선된 E-MTN을 사용했을 때 성능이 더욱 향상됨을 확인하였다. 향후 연구에서는 다중 양식을 더욱 효율적으로 사용할 수 있는 융합 방식과 모델 구조의 개발을 통해 비디오 폭력 탐지 시스템의 정확성과 효율성을 더욱 높이는 방향으로 진행할 예정이다.

Acknowledgment

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2019-0-01183)의 지원을 받아 작성됨.

참고 문헌

[1] W. Sultani, C. Chen, M. Shah, "Real-world anomaly detection in surveillance videos," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6479-6488.

[2] W. Chen, K. T. Ma, Z. J. Yew, M. Hur, D. A. A. Khoo, "TEVAD: Improved video anomaly detection with captions," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 5549-5559.

[3] P. Wu, J. Liu, Y. Shi, Y. Sun, F. Shao, Z. Wu, Z. Yang, "Not only look, but also listen: Learning multimodal violence detection under weak supervision," Computer Vision - ECCV 2020: 16th European Conference, Springer International Publishing, Part XXX 16, Glasgow, UK, August 23 - 28, 2020, pp. 322-339.

[4] Y. Tian, G. Pang, Y. Chen, R. Singh, J.W. Verjans, G. Carneiro, "Weakly-Supervised Video Anomaly Detection with Robust Temporal Feature Magnitude Learning," Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 4975-4986.

[5] C. Zhang, G. Li, Y. Qi, S. Wang, L. Qing, Q. Huang, M. H. Yang, "Exploiting Completeness and Uncertainty of Pseudo Labels for Weakly Supervised Video Anomaly Detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 16271-16280.

[6] H. Zhou, J. Yu, W. Yang, "Dual Memory Units with Uncertainty Regulation for Weakly Supervised Video Anomaly Detection," Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 37, No. 3, 2023, pp. 3769-3777.

차량 문콕사고 UX개선을 위한 연구방향 탐색

윤경빈¹, 윤봉식²

남부대학교 일반대학원 기계공학석사과정¹, 남부대학교 자동차기계공학과²
 rmfpdl28@nambu.ac.kr¹, jscomi@nambu.ac.kr²

A Study Direction for UX Improvement in Vehicle Door-ding Accident

Gyeong Bin Yun¹, Bong-Shik Yun²

Master's Course, Department of Mechanical Engineering, Nambu Univ¹
 Department of Automotive & Mechanical Engineering, Nambu University²

요약

Door dings frequently occur in daily life, leading to various issues such as vehicle damage, environmental concerns arising from repair processes, economic burdens due to increased insurance premiums, and conflicts among neighbors. The solutions derived from this study encompass improvements in public policy and functional solutions integrated within vehicles. Future research will focus on experimental studies to validate the three proposed directions. Preventing such issues can contribute to reducing societal costs and advancing ESG objectives.

1. 서론

차량의 문을 열 때 주변 차량에 문을 부딪쳐 파손을 입히는 행위를 문콕이라 한다. 문콕은 차량 승하차 간 가장 많이 발생하는 사고이며 여러 가지 사회·경제적 문제를 야기한다. 환경적으로는 차량의 기스나 패임을 해결하기 위해 사용되는 공정 간 오염원, 경제적으로는 보험비용 상승과 피해자의 기회비용 문제, 가장 중요한 대인 간 다툼으로 인한 사람 간의 분쟁과 신뢰성 저하 등 다양한 문제로 이어지고 있다[1].

이러한 문제가 발생되지 않도록 교육적 개도 등 대책도 필요하나 본 연구에서는 직접적인 예방차원의 접근을 시도하고자 한다.

현재까지 일부 차량의 스펙에 도어의 개폐 방식변화, 도어 센서 등 기술적 접근이 시도되었으나, 본질적으로 주차장의 주차구획을 확장시키거나[2] 자동차 제조사의 기능적 개선을 통한 근본적인 대책은 미비한 현황이다. 물론 가장 중요한 건 차량탑승자들의 세심한 주의와 배려하는 운전 습관이 필요하나 주차 시 경미한 차량파손이나 불법주정차에 관한 스마트블린신호 민원은 양적으로도 큰 폭으로 상승 중이다.[3]

이에 본 연구에서는 운전자의 스트레스요인과 요인별 스트레스지수, 전문가 탐문을 통해 문콕으로 인한 문제를 줄이는데 도움이 될 수 있는 사회적관점, 자동차기술관점을 중심으로 연구방향 탐색을 진행하였다.

2. 관련 연구

2.1 운전자 스트레스 요인 조사

광주지역 운전자를 대상으로 2022년 8월~10월 간 조사한

스트레스요인 조사에 유효 응답자 237명 답변을 분석한 결과 승차, 하차, 승하차 복합, 주차 간 스트레스 상위 요인은 주차공간문제와 모두 연결되어 있고, 복수 응답자 중 69%이상이 문콕사고와 관련된 스트레스를 경험한 것으로 조사되었다.



(그림 1) 승하차 및 주차 시 스트레스 요인 조사

요인별 스트레스 수준에 관한 조사에서 주차공간과 관련된 요인들과 문콕 등 차량에 발생하거나 발생이 예상되는 요인들에 5점 척도 내 4점 이상의 높은 반응을 보였다.



(그림 2) 운전자 스트레스의 요인별 지수 조사

2.2 국내 문콕관련 사고 빈도

2024년 8월 조사한 광주광역시 소재 자동차정비업소 대상 월평균 사고 차량 중 문콕 사고 비율에 대한 인터뷰 조사 결과 수리차량이 가장 많은 영업소의 경우 월평균 150건 중 100건을 초과하며, 수리차량이 적은 영업소의 경우 70건 중 60건을 상회한다고 조사되었다.

2.3 국내 차량 규모 및 주차장 기준

카이즈유 데이터연구소에 따르면 중형 이상 부피가 큰 차량은 57.3%인 695만대로 좁은 주차장에서 발생할 사고와 민원의 원인이 크게 증가한 것으로 보고되었다[4]. 2019년 개정된 주차장법 시행규칙 제3조에 따르면 주차장의 주차구획은 표 1과 같이 단위구획은 흰색 실선을 사용하며 경형자동차 전용주차구획은 파란색 실선으로 표시하도록 되어있다. 또한 둘 이상의 연속된 주차단위구획의 총 너비 또는 총 길이는 제1항에 따른 주차단위구획의 너비 또는 길이에 주차단위구획의 개수를 곱한 것 이상이 되어야 한다.

(표 1) 평행주차형식의 경우

구분	너비	길이
경형	1.7m 이상	4.5m 이상
일반형	2.0m 이상	6.0m 이상
보도와 차도 구분 없는 주거지역 도로	2.0m 이상	5.0m 이상
이륜자동차전용	1.0m 이상	2.3m 이상

(표 2) 평행주차형식 외의 경우

구분	너비	길이
경형	2.0m 이상	3.6m 이상
일반형	2.5m 이상	5.0m 이상
확장형	2.6m 이상	5.2m 이상
장애인전용	3.3m 이상	5.0m 이상
이륜자동차 전용	1.0m 이상	2.3m 이상

이는 일반형 주차장 너비 최소 기준이 20cm, 확장형 주차장 너비와 길이가 10cm씩 넓어진 것으로 이전 국내 규격인 2.3×5.0 m는 일본의 소형차 주차구획과 비슷했고 개정 후 기준인 2.5×5.0m는 미국의 2.7×5.5m, 일본 개정기준인 2.5×6.0m와 비교된다.

2.4 문콕사고 처리 방식

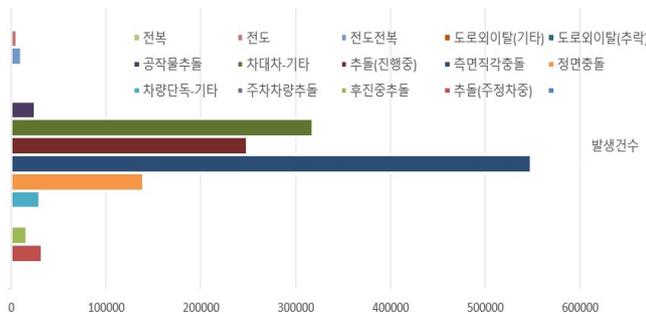


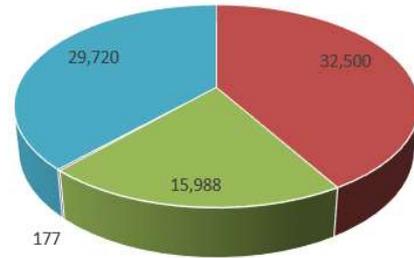
그림 3 전체 사고유형별 사고 빈도

현재 문콕사고는 보험처리나 합의금 방식으로 처리되며 문콕 가해차량의 보험 여부에 따라 당사자 간 보험처리로 협의가 되면 보험사에 접수와 보상절차를 진행한다. 다만

문콕 보험처리로 인한 보험료 할증요인으로 수리에 따른 합의금 선에서 사건이 마무리되는 경우도 많다고 한다.

2018년 카이즈유 매거진에 따르면 문콕 사고관련 보험 청구 건수는 2014년 2,200건, 2015년 2,600건, 2016년 3,400건에 달했고 매년 증가세를 보였고, 공식적인 건수 외에 실제 발생 건수는 훨씬 더 많을 것으로 추측된다.

공공데이터포털의 2015~2022년 간 교통사고 통계자료 기준 사고유형별 사고빈도는 그림3과 같이 다양한 유형으로 분류가 가능하며 이 중 본 연구와 관련된 주차와 관련된 사고유형은 11.9%에 이른다.



■ 추돌(주정차중) ■ 후진중추돌 ■ 주차차량추돌 ■ 차량단독-기타

그림 4 주차 중 사고유형별 사고 빈도

3. 결론 및 후속연구 방향

문콕과 관련한 사전 연구 결과 연구에 대한 가설단계의 예측보다 많이 상회한 수의 관련 사고가 발생되고 있는 것을 알 수 있었다. 또한 주차 공간에 대한 국민의 스트레스와 문콕사고로 이어진 사회적 경제적 문제 및 사람 간의 신뢰문제나 폭언과 폭행 등 2차 피해가 빈번히 발생하고 있어 국가적 교육과 제도 외에도 직접적인 예방책이 요구되고 있다.

이에 주차장에 관한 법령 개정 이전 문콕 발생률 대비 개정 후 발생률과 해외의 문콕 발생률이 주차공간에 미치는 영향에 대한 연구를 통해 현재 운용 중인 모빌리티환경에 적합한 주차공간 연구와 기능적 예방 방법에 대한 후속 연구를 진행하고자 한다.

연구성과를 적용한 기술적 목표는 센서의 음성 및 물리적 정보와 주차공간 인식 기술[5], 기능적 회피 방법에 주목하고 있다.

참고 문헌

[1] 임상현 외, 경미사고 사례분석을 통한 공학적 상해위험 분석의 유효성, 한국법과학회지 24[1], 53~57pp, 2023
 [2] 김혜진 외, PM교통사고에 영향을 미치는 근린환경 요인분석: 로지스틱 회귀모형을 이용하여, 대한교통학회지, 40[2], pp.205~217, 2022
 [3] 박준상 외, 불법 주차차에 영향을 미치는 도시환경 요인 분석:서울시 스마트불편신고 민원자료를 중심으로, 지역연구, 38[3], pp.3~17, 2022
 [4] 카이즈유데이터연구소 매거진 6월호, 2024.07.25.
 [5] <http://journal.ksae.org/xml/41344/41344.pdf>

아이템 영역 마스크를 이용한 딥러닝 기반 패션 아이템의 무늬 패턴 분류

권영채, 최진영, 김인중
 한동대학교 전산전자공학과

e-mail : {youngchaekwon, jinyoung}@handong.ac.kr, ijkim@handong.edu

Deep Learning-based Fashion Item Pattern Classification Using Item-Region Mask

Youngchae Kwon, Jinyoung Choi, and Injung Kim

Dept of CSEE, General Graduate School, Handong Global University

요약

전자 상거래 시장 규모가 성장함에 따라 옷, 모자, 신발 등 다양한 패션 아이템의 온라인 구매 역시 빠르게 확대되고 있다. 이에 따라 사용자가 원하는 패션 아이템의 특성을 이해하고 추천해 주는 시스템이 부각되고 있다. 패션 아이템은 색상, 무늬 패턴, 길이 등 다양하고 고유한 특성을 가질 수 있으며 각 특성에 대해 정확한 분석이 필요하다. 본 연구는 아이템 영역 마스크를 활용해 18개의 패션 아이템 무늬 패턴을 분류하는 방법을 제안하고 실험을 통해 제안하는 네트워크의 효과를 입증하였다.

1. 서론

전자 상거래 시장의 발전과 함께 의상, 액세서리 등 패션 아이템의 온라인 구매 역시 증가하고 있다. 소비자의 구매를 유도하기 위해서는 소비자가 선호하는 속성의 패션 아이템을 추천해야 한다. 이에 따라 패션 아이템의 색상, 패턴, 소재 등의 속성을 분석해야 한다. 본 연구에서는 패션 아이템의 무늬 패턴을 인식하는 방법을 제안한다.

제안하는 모델은 CLIP[1] 기반 영상 분할 모델 CLIPSeg[2]를 통해 패션 아이템 영역의 제로 샷(zero-shot) 마스크를 추출하고, 이를 활용해 패션 아이템의 무늬 패턴을 분류한다. CLIPSeg가 추정된 아이템 영역에 집중해 패턴을 인식함으로써 분류 성능을 개선했으며, 상의, 겉옷, 하의 등 의류와 가방, 안경 등 패션 소품을 포함한 다양한 실생활 데이터에 범용적으로 적용될 수 있다.

2. 관련 연구

2.1 CLIPSeg

CLIP(Contrastive Language-Image Pretraining)[1]은 대규모 이미지와 텍스트 쌍의 대조 학습을 통해 이미지와 텍스트 간 호환이 가능한 공동 임베딩 공간을 학습하는 모델이다. 이를 기반으로 제안된 영상 분할 모델 CLIPSeg[2]는 쿼리 이미지와 함께 프롬프트를 입력받아 이미지 내에서 프롬프트에 의해 지정된 영역에 대한 분할 마스크를 출력한다. 특히 CLIPSeg는 대규모 사전학습을 통해 우수한 제로 샷 예측 성능을 갖는다.

2.2 Swin Transformer

Swin Transformer[3]는 이동창(shifted window)을 이용해 ViT(Vision Transformers)[4]를 계층 구조로 확장한

모델이다. 다양한 크기의 이미지 패치에 대한 벡터 표현을 학습하기 때문에 다양한 크기의 개체를 다룰 수 있으며, 이동창을 통해 ViT의 시간 복잡도를 선형 복잡도로 줄였다. Swin Transformer는 이미지 분류, 객체 탐지, 이미지 분할 등 컴퓨터 비전 분야의 여러 작업에서 백본(backbone) 네트워크로 사용할 수 있도록 설계되었다.

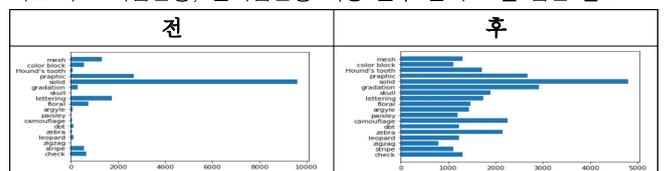
3. 제안하는 방법

3.1 데이터 수집 및 전처리

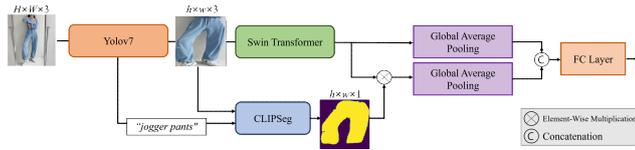
본 연구는 패션 플랫폼 기업 Deep Fashion이 수집한 Showniq 데이터셋을 학습과 평가에 사용하였다. Showniq 데이터셋은 30,121개의 패션 아이템을 포함한 27,799장의 이미지로 구성되어 있다. 개별 패션 아이템에 대해 바운딩 박스와 무늬 패턴 클래스 레이블을 포함한다. 무늬 패턴 클래스로는 check, stripe, zigzag, leopard, zebra, dot, camouflage, paisley, argyle, floral, lettering, skull, gradation, solid, graphic, hound's tooth, color block, mesh 등 총 18종이 포함되었다.

Showniq 데이터셋은 표 1 좌측 그래프와 같이 무늬 패턴 클래스 별 샘플 수의 편차가 크다. 본 연구에서는 이를 극복하기 위해 표 1 우측 그래프와 같이 오버샘플링과 언더샘플링을 적용해 샘플 수 편차를 줄였다.

(표 1) 오버샘플링, 언더샘플링 적용 전후 클래스 별 샘플 분포



3.2 모델 설계



(그림 1) 제안하는 네트워크

제안하는 모델의 구조는 그림 1과 같다. 먼저 패션 아이템이 포함된 RGB 이미지를 입력받고, 별도로 학습된 YOLOv7[5]을 이용해 패션 아이템의 영역과 종류를 예측한다. 예측된 아이템 영역에 따라 아이템 단위로 이미지를 크롭하고 Swin Transformer를 이용해 크롭된 이미지의 특징 맵을 추출한다. 한편, CLIPSeg의 쿼리 이미지와 텍스트 프롬프트로써 각각 크롭된 RGB 이미지와 YOLOv7을 통해 예측된 아이템의 종류를 입력받아 아이템 영역 마스크를 추출한다. 추출된 아이템 영역 마스크는 Otsu 이진화와 닫기 연산(closing operation)을 통해 이진 아이템 영역 마스크로 변환한다. 이진 아이템 영역 마스크는 Swin Transformer를 통해 추출한 이미지 특징 맵과의 원소 간 곱셈 연산을 통해 아이템 특화 특징 맵을 추출하는 데 사용된다. 분류기는 이미지의 특징 맵과 아이템 특화 특징 맵을 결합하고 결합된 특징 맵에서 FC(fully-connected) 계층을 통해 패션 아이템 무늬 패턴을 분류한다.

4. 성능 평가

4.1 실험 세팅

모델은 사전 학습된 Swin Transformer-Base를 전이해 사용했으며, SGD를 이용해 ShowNiq 데이터에 미세조정했다. 초기 학습률은 0.00001로 설정했으며 학습 스케줄링은 LambdaLR 전략을 사용했다. 손실함수로 크로스 엔트로피를 적용했다.

4.2 실험 결과

표 2와 같이 제안하는 네트워크는 18개의 무늬 패턴 클래스에 대해 top-1 정확도 78.46%의 성능을 달성하였다. 이는 이진 아이템 영역 마스크를 활용하지 않는 Swin Transformer[3]의 68.65%에 비해 9.81% 개선된 결과이다.

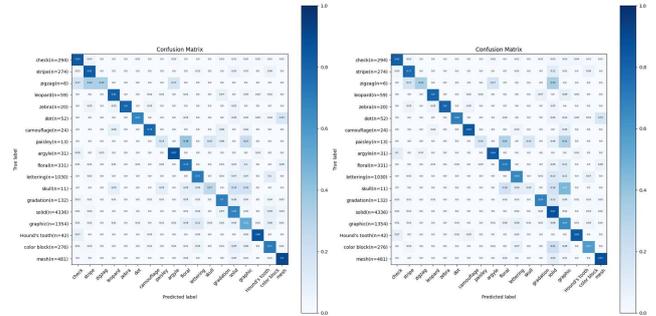
그림 2와 같이 zigzag, paisley, skull 등 샘플 수가 적은 클래스는 check, solid, mesh 등 샘플 수가 많은 클래스와 비교했을 때 상당히 낮은 정확도를 보였다. 이는 샘플 수가 적은 클래스의 데이터가 보완될 경우 성능이 개선될 수 있는 가능성을 시사한다.

(표 2) 패션 아이템 무늬 패턴 분류 정확도 (%)

	top-1	top-2	top-3	top-4	top-5
Swin Transformer[3]	68.65	87.51	93.50	96.78	98.17
Ours	78.46	91.41	95.79	98.00	98.86

5. 결론

본 연구에서는 의류 및 패션 소품을 포함한 다양한 패션 아이템 이미지의 무늬 패턴을 분류하는 모델을 제안했다. CLIPSeg를 이용해 추출한 이진 아이템 영역 마스크를 통해 아이템 영역에 집중해 무늬 패턴을 인식하였다. 실험에서 제안하는 방법은 Swin Transformer와 비교해 9.81% 개선된 성능을 보여 제안하는 방법이 효과적임을 입증하였다.



(그림 2) Confusion Matrix (Swin (left), Ours (right))

사사의 글

본 연구는 ㈜딥패션과 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2023-0-00055)

참고 문헌

- [1] Alec Radford et al., "Learning Transferable Visual Models from Natural Language Supervision", International Conference on Machine Learning. PMLR, 2021. pp.8748-8763.
- [2] Timo Lüddecke, Alexander Ecker, "Image Segmentation Using Text and Image Prompts", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 2022. pp.7086-7096.
- [3] Ze Liu et al., "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows", Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. pp.10012-10022.
- [4] Alexey Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", Proceedings of the International Conference on Learning Representations. 2020.
- [5] Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan M. Liao, "Yolov7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023. pp.7464-7475.

감자 잎 질병 감지를 위한 ViT 기반 딥 러닝 모델

Md Ilias Bappi*
전남대학교
*_bappi@jnu.ac.kr

김경백**
전남대학교

ViT-Based Deep Learning Model for Potato Leaf Disease Detection

Md Ilias Bappi*
Dept. of. Artificial Intelligence
Convergence, Chonnam National
University

Kyungbeak Kim**
Dept. of. Artificial Intelligence
Convergence, Chonnam National
University

Abstract

This study aims to develop a cutting-edge automated system for the early detection and classification of leaf diseases in potato plants, one of the most widely cultivated crops globally. Diseases like early and late blight, caused by *Alternaria solani* and *Phytophthora infestans*, significantly impact potato production. To address this, we propose a model based on the Vision Transformer (ViT) with Multi-Head Attention (MHA) to accurately identify these diseases, overcoming the limitations of traditional methods that are often labor-intensive, time-consuming, and prone to inaccuracies. The model's performance is evaluated using Precision, Recall, Accuracy, and F1 score, achieving an accuracy of 99.83%. These results suggest that Vision Transformers offer a powerful approach for plant disease detection in agricultural research.

Keywords – Vison Transformer, Multi Head Attention, Potato Leaf Classification, Deep learning.

1. Introduction

The Food and Agriculture Organization of the United Nations (FAO) predicts a global population of 9.1 billion by 2050, raising food demand [1]. However, agricultural growth and water availability lag behind, with crop diseases posing major threats, particularly in developing nations. Traditional plant evaluations are labor-intensive and unreliable [2], making rapid, accurate disease detection crucial. Researchers are thus turning to automated systems to address these challenges [3]. Potatoes, the third most consumed staple food globally after rice and wheat, feed over a billion people. However, diseases originating from potato leaves cause a 9-11% annual reduction in output, hindering crop growth [4]. While initial studies relied on complex and costly biological techniques [5], image processing and machine learning (ML) methods have emerged as promising alternatives for crop disease surveillance. Traditional ML techniques like Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) have been widely used for plant disease classification [14]. However, deep learning (DL) approaches, particularly Vision Transformer (ViT), have proven more effective by leveraging global attention models. Overall, our proposed model addresses the challenges in image classification by leveraging the strengths Transformer, and attention architectures, resulting in an effective and efficient solution for disease classification in potato leaves. In order to categorize multi-class diseases of potatoes, this research suggests a better

method, ViT, based on MHA.

2. Related works

In the field of potato leaf disease detection from images, various approaches have been explored. Convolutional Neural Networks (CNNs), such as LeNet, AlexNet, VGGNet, GoogLeNet, ResNet, and DenseNet, have significantly advanced the field. CNNs have been used in agriculture for tasks like monitoring olive tree responses and detecting live blueberry plants [6]. Recently, Transformers, known for their success in Natural Language Processing (NLP) with models like BERT and GPT, have been adapted for image processing. The Vision Transformer (ViT) applies transformer models directly to images, offering advantages over CNNs by providing a comprehensive global view and better representation similarity across different levels of information. The skip connection in ViT further enhances its performance. This study focuses on using ViT, including its attention mechanisms, to improve potato disease identification.

3. Methodology

For this study, we used a dataset of 1,000 images each of early blight, late blight, and healthy potato leaves from kaggle [7]. We employed the Vision Transformer (ViT) with Multi-Head Attention (MHA) for classification, a state-of-the-art architecture that excels in object classification. Unlike traditional Convolutional Neural Networks (CNNs), which rely on convolutions,

the ViT treats an image as a sequence of patches and processes it using a Transformer encoder, similar to those used in Natural Language Processing. An overview of the ViT architecture is illustrated in Fig 1. The input images were resized to 224×224 pixels. We

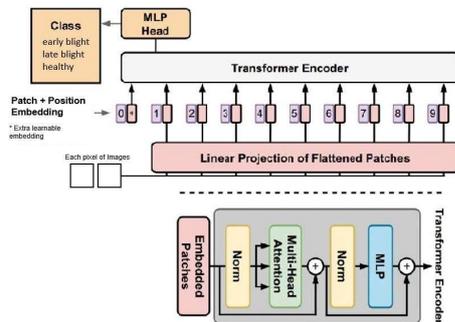


Fig1: Vision Transformer architecture – main blocks.

used a pre-trained ViT model, which was further fine-tuned with attention mechanism architectures for 50 epochs. The fine-tuning was conducted using the PyTorch framework on a system with 64GB RAM and two GPUs (GForX 2). The network was optimized with Stochastic Gradient Descent (SGD) and a momentum of 0.9. The initial learning rate (LR) was set to 0.01 and adjusted adaptively: if the validation loss did not improve for two consecutive epochs, the learning rate was reduced by 10%. We maintained a mini-batch size of 32 throughout the training. Our approach achieved an impressive accuracy of 99.83%

4. Result

The proposed method achieved outstanding performance in classifying potato leaf diseases, reaching an accuracy of 99.83%. It also attained an F1 Score, Precision, and Recall of 0.9983, demonstrating a high balance between precision and recall. The model was fine-tuned with a mini-batch size of 32 for 50 epochs, using an adaptive learning rate. The training loss was 0.0060, and the validation loss was 0.0037, indicating excellent generalization with minimal overfitting (see Fig2). These results underscore the model’s robustness and effectiveness in accurately identifying potato leaf diseases.

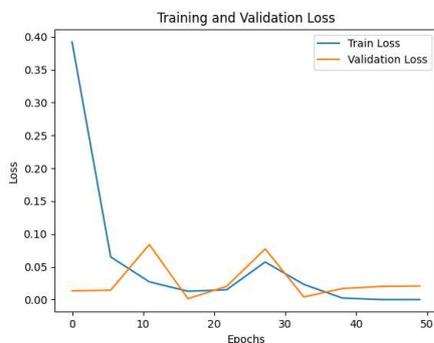


Fig2: Train and validation loss curve.

4.1 Comparison with others methods

The proposed model outperforms other architectures with an accuracy of 99.83%, significantly higher than CNN (88.65%), MLP-Mixer (82.77%), and Swin

Transformer (90.31%). It also leads in sensitivity (95.13%) and precision (95.71%), compared to CNN (89.52% and 86.92%), MLP-Mixer (74.29% and 90.10%), and Swin Transformer (81.87% and 60.42%). Its F1 Score of 97.82% surpasses all, demonstrating its superior classification performance (See Table 1).

Architecture	ACC	SEN	SPE	PRE	F1SCO
CNN	88.65	0.8952	0.4375	0.8692	0.882
MLP-Mixer	82.77	0.7429	0.759	0.901	0.8143
Swin Transformer	0.9031	0.8187	0.0887	0.6042	0.6704
Proposed model	0.9983	0.9513	0.8097	0.9571	0.9782

Table 1: Comparison with others recent models.

5. Conclusion

In this study, we developed a Vision Transformer (ViT) model with Multi-Head Attention (MHA) for classifying potato leaf diseases, achieving an accuracy of 99.83% and outperforming traditional models like CNNs, MLP-Mixer, and Swin Transformer. The model demonstrated strong generalization with low training and validation losses, making it highly effective for agricultural applications. For future work, we plan to apply transfer learning to enhance performance and efficiency further. Additionally, we will test the model on larger datasets to fully exploit the potential of transformer architectures, ensuring scalability and broader applicability in agricultural contexts.

Acknowledgement

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry(IPET) through the Agriculture and Food Convergence Technologies Program for Research Manpower development, funded by Ministry of Agriculture, Food and Rural Affairs(MAFRA)(project no. RS-2024-00397026). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government(MSIT).

References

- [1] Bruinsma, J. The Resource Outlook to 2050: By How Much Do Land, Water and Crop Yields Need to Increase by 2050; Food and Agriculture Organization of the United Nations: Rome, Italy, 2009; pp. 24 - 26.
- [2] Pantazi, X.; Moshou, D.; Tamouridou, A. Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. *Comput. Electron. Agric.* 2019, 156, 96 - 104. [CrossRef]
- [3] Wolfenson, K.D.M. Coping with the Food and Agriculture Challenge: Smallholders’ Agenda; Food Agriculture Organization of the United Nations: Rome, Italy, 2013
- [4] Sardogan, M.; Tuncer, A.; Ozen, Y. Plant leaf disease detection and classification based on CNN with LVQ algorithm. In Proceedings of the 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia and Herzegovina, 20 - 23 September 2018; pp. 382 - 385.
- [5] Sankaran, S.; Mishra, A.; Ehsani, R.; Davis, C. A Efficient RMT-Netew of advanced techniques for detecting plant diseases. *Comput. Electron. Agric.* 2010, 72, 1 - 13.
- [6] Hellmann, H.; Goyer, A.; Navarre, D.A. Antioxidants in potatoes: A functional view on one of the major food crops worldwide. *Molecules* 2021, 26, 2446.
- [7] kaggle, PLD, <https://www.kaggle.com/datasets/muhammadardiputra/potato-leaf-disease-dataset>.

ResNet-50와 U-Net을 이용한 다중 클래스 치과 영상 분할

아프난 가푸어, 이범식*
정보통신공학과, 조선대학교
{afnanghafoor, bslee*}@chosun.ac.kr

Multi-Class Dental Image Segmentation Using ResNet-50 and U-Net

Afnan Ghafoor, Bumshik Lee*
Department of Information and Communication Engineering, Chosun
University

Abstract

Dental image segmentation plays a important role in accurate and efficient dental diagnostics. In this paper, we propose a segmentation approach that integrates a ResNet-50 encoder with a U-Net-based decoder. This method is designed for multi-class segmentation, demonstrating excellent performance in distinguishing between 32 different dental classes. The high precision in identifying dental features is due to ResNet-50's deep architecture, allowing the model to capture even the smallest details. Additionally, U-Net's spatial localization capabilities in the decoder ensure precise segmentation of each tooth and its related structures. Our experimental results indicate that the proposed method surpasses current state-of-the-art techniques, delivering superior accuracy across various segmentation metrics.

1. Introduction

Dental imaging is essential for teeth diagnosis, with panoramic X-rays providing a comprehensive view of the entire mouth, including teeth, jaws, and surrounding structures, unlike intraoral X-rays that focus on individual teeth. This broader perspective aids in detecting dental issues like cavities, bone abnormalities, cysts, tumors, and impactions. However, analyzing panoramic X-rays presents challenges due to their complexity and variability, often leading to suboptimal classification performance in traditional diagnostic methods. Manual interpretation can be time-consuming, prone to variability, and risk overlooking minor abnormalities.

The advent of machine learning, particularly deep learning, has transformed medical imaging. Convolutional neural networks (CNNs)[2] have demonstrated excellent image recognition capabilities, paving the way for automated, precise, and efficient diagnostics. This study leverages these advancements to address the segmentation of panoramic X-rays. Rather than limiting the scope to binary or multi-class problems, we expand it to multi-class segmentation with 32-class categorization. We propose a model for complex pixel classification that enhances diagnostic accuracy and efficiency in dental care.

Our model incorporates a pre-trained ResNet-50 encoder, known for its robust feature extraction, coupled with a U-Net decoder, recognized for its localization effectiveness in biomedical segmentation tasks. Skip connections are utilized to

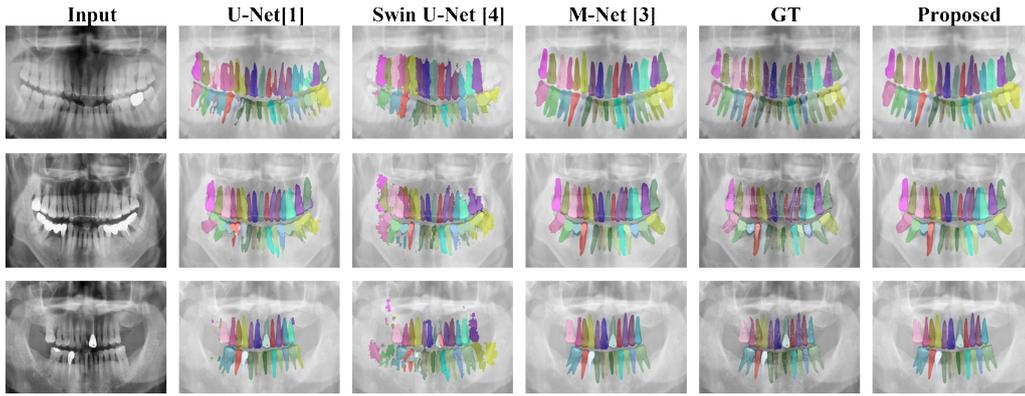
retain best features during encoding. This hybrid ResNet-50 and U-Net[1] architecture excels in segmenting panoramic X-rays, offering superior performance compared to other methods.

2. Proposed methodology

This paper uses a ResNet-50 encoder, selected for its ability to extract both high-level and fine-grained features, such as tooth shapes and textures. The U-Net-based decoder, known for its accuracy in object localization, is crucial for segmenting complex dental structures. By combining ResNet-50's feature extraction capabilities with U-Net's precision, our model addresses challenges like inconsistent structure identification and boundary definition inaccuracies. This architecture is specifically tailored for dental images, ensuring high segmentation accuracy and computational efficiency.

The ResNet-50 encoder, pre-trained on the ImageNet1k dataset, captures comprehensive features from panoramic dental X-rays. Its deep structure allows for detailed feature extraction, detecting subtle irregularities and boundary structures more effectively than traditional encoders.

The U-Net-like decoder, designed to tackle the complexity of dental images, accurately localizes structures while minimizing false positives and maintaining sharp boundaries. This combination of encoder and decoder provides superior segmentation accuracy compared to existing models.



Picture 1. Comparison between the proposed and the conventional methods

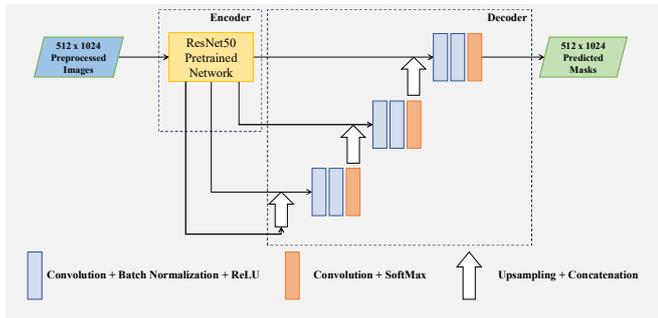


Figure 1. Overall architecture of the proposed network

The model is trained using a loss function that describes as below:

$$L = 1 - \frac{(2\sum(y_t \cdot y_p)^2 + \epsilon)}{(2\sum(y_t^2 + y_p^2) + \epsilon)}$$

where L is the loss function, y_t and y_p represent the ground truth and predicted segmentation maps, respectively, and ϵ is a smoothing factor for avoiding division by zero.

2. Experimental results

The proposed model was evaluated against U-Net [1], M-Net, and Swin-Unet [3] using DSC, Jaccard Index (JI), Precision, and Recall metrics as shown in Table 1. Our model achieved the highest scores across all metrics, with a DSC of 0.8012, JI of 0.7351, Precision of 0.7602, and Recall of 0.8655. This superior performance demonstrates the model's effectiveness in accurately segmenting dental structures, surpassing existing state-of-the-art methods. Notably, the higher Recall score reflects the model's ability to capture even the finer details in dental X-ray images, reducing the likelihood of missed detections. Additionally, the combination of deep feature extraction from ResNet-50 and the precise localization capabilities of the U-Net decoder enables the model to handle complex structures more efficiently. These results suggest that the proposed model holds promise for improving the accuracy and reliability of dental diagnostics

Table 1. Quantitative comparisons of the methods

Models	DSC	JI	Precision	Recall
U-Net [1]	0.7602	0.6871	0.7458	0.8366
M-Net	0.7846	0.7132	0.7557	0.8391
Swin-Unet [3]	0.6348	0.5296	0.6107	0.7192
Proposed	0.8012	0.7351	0.7602	0.8655

Conclusion

This paper introduces a method for dental X-ray segmentation, integrating the powerful feature extraction of ResNet-50 with the precise localization of a U-Net-based decoder. Our model efficiently captures both broad patterns and fine details of dental structures in panoramic images. Results show that it consistently outperforms current state-of-the-art techniques in both quantitative metrics and visual evaluations.

Acknowledgement

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE) (2021RIS-002). We appreciate the high-performance GPU computing support of HPC-AI Open Infrastructure via GIST SCENT

References

- [1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015, arXiv:1505.04597.
- [2] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in Proc. IEEE Int. Conf. Comput. Vis., Venice, Italy, Jun. 2017, pp. 2961-2969.
- [3] H. Cao et al., 'Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation', arXiv [eess.IV]. 2021.

크리에이터 산업의 지속 가능성: 위기와 기회 (MCN 기업을 중심으로)

권병민, 이지은
한양사이버대학교 경영대학원 ITMBA
e-mail : benjamin@urbanomad.xyz, scully1215@hycu.ac.kr

The Sustainability of the Creator Industry: Challenges and Opportunities (Focusing on MCN Companies)

Byungmin Kwon, Ji-Eun Lee
Hanyang Cyber University, Graduate School of Business, ITMBA

요약

본 연구는 유튜브, 틱톡 등 뉴미디어 플랫폼의 확산과 함께 급성장한 크리에이터 산업과 이를 지원하는 MCN(Multi-Channel Network) 기업의 역할과 도전 과제를 분석한다. 크리에이터는 오늘날 대중 문화의 중요한 축으로 부상하며, 다양한 콘텐츠를 통해 사회적 영향력을 확장하고 있다. 이에 따라 MCN 기업은 크리에이터의 콘텐츠 유통, 저작권 관리, 광고 유치 등에서 중요한 역할을 수행하며, 광고 수익, 상품 판매, IP 부가 사업 등 다양한 비즈니스 모델을 개발하고 있다. 그러나, MCN 기업들은 지속 가능한 경영 모델을 확립하는 데 어려움을 겪고 있으며, 크리에이터 이탈 등 위험 요소도 존재한다. 특히, 많은 MCN 기업들이 수익성을 확보하지 못하고 있으며, 적자 경영을 지속하는 상황이다. 본 연구는 이러한 문제를 해결하기 위한 전략으로 MCN 기업의 사업화 역량 강화, 크리에이터의 체계적인 기획 및 육성, 레이블 체계 도입을 제안한다. 이러한 전략은 MCN 기업과 크리에이터 간의 협력을 강화하고, 더 높은 수익 배분율을 통해 MCN 기업과 크리에이터 산업의 지속 가능성을 높이는 데 이바지할 수 있다. 본 연구는 이론적으로 크리에이터 산업과 MCN 기업 간의 상호작용을 이해하는 데 도움을 주며, 실무적으로는 지속 가능한 비즈니스 모델 구축을 위한 전략적 방향을 제시한다.

주제어: 지속 가능성, 비즈니스 모델, 크리에이터 산업, MCN

1. 서론

크리에이터 산업은 유튜브, 틱톡 등 뉴미디어 플랫폼의 확산과 함께 급속도로 성장하고 있다. 크리에이터는 이제 단순한 취미를 넘어 대중 문화의 중심이 되었으며, 초등학교생들의 장래 희망 직업으로도 손꼽히고 있다[1]. 이러한 변화는 크리에이터가 창출하는 사회적 영향력이 점점 커지고 있음을 반영한다. 이와 동시에, 크리에이터의 활동을 지원하고 관리하는 MCN(Multi-Channel Network) 기업들도 산업의 중요한 축으로 자리 잡았다. 그러나, MCN 기업들은 수익성 부족과 지속 가능한 경영 모델의 부재 등 여러 도전 과제에 직면해 있다. 이러한 맥락에서 본 연구는 크리에이터 산업과 MCN 기업의 현재 상황을 분석하고, 지속 가능한 성장을 위한 전략을 모색하고자 한다.

2. 관련 연구

2.1 크리에이터 산업의 발전과 MCN 기업의 등장

크리에이터는 디지털 플랫폼에서 텍스트, 이미지, 동영상 등 다양한 형식의 콘텐츠를 제작하고 유통하는 사람을 의미한다. 이들은 유튜버, 틱톡커, 인스타그램어, 스트리머 등으로 불리며, 각 플랫폼에서 독자적인 콘텐츠와 영향력을 발휘하고 있다. 크리에이터 산업의 성장과 함께, 이들을 지원하고 관리하는 MCN 기업이 등장했다. MCN은 여

러 크리에이터를 대표하며, 콘텐츠 유통, 판매, 저작권 관리, 광고 유치 등의 업무를 수행한다. 2013년 CJ ENM이 국내 최초로 MCN 사업을 시작한 이후[2], 샌드박스네트워크, 트레저헌터 등 다양한 기업이 이 시장에 진입하여 빠르게 성장했다.

2.2 MCN 기업의 비즈니스 모델

MCN 기업은 주로 다섯 가지 비즈니스 모델을 기반으로 운영된다. 첫째, 플랫폼 수익은 유튜브와 같은 플랫폼에서 발생하는 광고 수익으로, MCN과 크리에이터가 이를 분배한다. 둘째, 광고 캠페인은 브랜드드 콘텐츠나 PPL 형태의 협업을 통해 수익을 창출하는 모델로, MCN은 적합한 크리에이터와의 협력을 통해 효과적인 광고를 제공한다. 셋째, 상품 판매는 공동구매나 라이브커머스와 같은 형태로 크리에이터의 구독자와 팬을 대상으로 상품을 판매하는 방식이다. 넷째, IP 부가 사업은 크리에이터의 인지도를 활용하여 굿즈, 팬미팅 등 부가적인 상품과 서비스를 개발하고 판매하는 것이다. 마지막으로, 콘텐츠 유통은 OTT나 TV 등 다양한 매체에 크리에이터 콘텐츠를 공급하는 것을 포함한다.

2.3 MCN 기업과 크리에이터의 협력 구조

MCN 기업과 크리에이터 간의 협력 관계는 크게 '전속

계약'과 '프로젝트 계약' 두 가지 형태로 나뉜다. 전속 계약은 MCN이 크리에이터의 모든 비즈니스 영역에서 독점적인 권리를 가지는 형태로, 크리에이터의 성장을 지원하고 다양한 비즈니스 기회를 제공한다. 반면, 프로젝트 계약은 특정 프로젝트에 한정하여 체결되며, 특정 캠페인이나 콘텐츠 제작에 필요한 경우 유연하게 적용된다. 이러한 계약 구조는 크리에이터의 성장 단계와 그에 따른 요구 사항에 따라 달라지며, 각 단계에서 크리에이터의 니즈에 맞는 지원을 제공한다.

2.4 MCN 기업의 도전 과제

MCN 기업들은 크리에이터 산업의 중심적인 역할을 수행하면서도 여러 도전 과제에 직면하고 있다. 첫째, 지속 가능한 비즈니스 모델을 구축하는 것이 어려운 문제로 남아 있다. 많은 대형 MCN 기업들이 여전히 수익을 내지 못하고 있으며, 이는 장기적인 안정성을 확보하는 데 큰 장애가 되고 있다. 표1과 같이, 최근 3년간 주요 대형 MCN 기업들의 실적을 보면, 대부분 적자를 기록하고 있으며, 이는 MCN 기업들의 재무적 압박을 가중시키고 있다. 둘째, 유명 크리에이터의 이탈 위험도 존재한다. 크리에이터와의 협력 관계를 유지하면서도 새로운 수익 모델을 개발해야 하는 MCN 기업들은 이 문제를 해결하기 위해 다양한 전략을 모색하고 있다.

(표 1) 주요 대형 MCN 기업 최근 3개년 실적 (단위: 억원)

회사명	항목	2021년	2022년	2023년
샌드박스	매출	1,137	1,462	1,028
	영업이익	-121	-212	-115
네트웍	매출	253	292	275
	영업이익	-54	-29	-36

3. 전략적 제언

MCN 기업이 이러한 도전 과제에 대응하기 위해서는 몇 가지 전략적 접근이 필요하다.

첫째, MCN 기업 주도적인 사업화 역량 강화가 중요하다. 플랫폼 수익이나 광고 캠페인과 같은 기존의 수익 모델뿐만 아니라, 상품 판매와 IP 부가 사업, 콘텐츠 유통과 같은 새로운 비즈니스 모델에 대한 역량을 강화해야 한다. 실제로 국내 대표 MCN 기업의 경우 안정적인 수익 창출을 위해 커머스 개발을 시도하고 있다[3]. MCN 기업이 주도적으로 기획하고 개발함으로써, 회사가 수익 배분에서 더 높은 비율을 확보할 수 있는 구조를 만들어 전반적인 수익성을 향상시킬 수 있다. 또한, 이 과정에서 크리에이터는 MCN 기업이 개발한 다양한 사업에 협력하게 되어, 양측 모두에게 이익이 되는 상생의 모델을 구축할 수 있다. 이를 통해, MCN 기업은 크리에이터와의 협력 관계를 강화하고, 더 높은 수익률을 달성함으로써 산업 전체의 지속 가능성을 높일 수 있다.

둘째, 크리에이터의 체계적인 기획과 육성이 필요하다. MCN 기업들은 크리에이터의 성장을 지원하고, 콘텐츠 기획 및 제작 역량을 강화하는 데 집중해야 한다. 이를 위해 크리에이터의 특성에 맞는 맞춤형 지원과 교육 프로그램을

을 제공하고, 크리에이터의 성장 단계에 따른 다양한 비즈니스 기회를 창출해야 한다. 이러한 접근은 K-Pop 아이돌 그룹의 육성 모델과 유사하다. 아이돌 그룹을 기획하고 훈련하는 과정에서 기획사는 각 멤버의 개성과 역량을 극대화하는 맞춤형 교육과 훈련 프로그램을 제공하며, 이를 통해 다양한 활동 기회를 창출한다. 아이돌 그룹의 성공을 위해 체계적으로 관리되고 지원받는 것처럼, 크리에이터도 콘텐츠 제작과 팬 관리, 브랜드 협업 등에서 성공하기 위해 체계적인 육성과 관리가 필요하다. MCN 기업은 크리에이터의 지속적인 성장을 도모하고, 이를 바탕으로 한 수익 다각화를 위해 아이돌 육성 전략을 벤치마킹하여 적용할 수 있다. 이러한 체계적인 접근은 크리에이터와의 장기적인 협력 관계를 강화하고, 더 큰 시장에서의 경쟁력을 높이는 데 이바지할 것이다.

마지막으로, 레이블 체계의 도입을 제안한다. 레이블 체계는 크리에이터의 창작과 경영 독립성을 최대한 보장하면서도, 성과에 따른 보상을 공유하는 방식으로, MCN 기업과 크리에이터 간의 신뢰를 강화할 수 있다. 이를 통해 MCN 기업은 핵심 크리에이터의 이탈을 방지하고, 장기적인 협력 관계를 유지할 수 있을 것이다.

4. 결론

크리에이터 산업은 뉴미디어 플랫폼의 확산과 함께 빠르게 성장하고 있으며, 이에 따라 MCN 기업의 역할과 중요성이 더욱 커지고 있다. 본 연구는 MCN 기업이 직면한 도전 과제를 분석하고, 이를 해결하기 위해 사업화 역량 강화, 크리에이터의 체계적인 기획 및 육성, 레이블 체계 도입의 세 가지 전략을 제안하였다. 이러한 전략은 크리에이터와 MCN 기업 간의 협력을 강화하고, 산업 전체의 지속 가능성을 높이는 데 중요한 역할을 할 수 있다.

이론적으로, 본 연구는 크리에이터 산업과 MCN 기업의 상호작용에 대한 심도 있는 이해를 제공하며, 실무적으로는 MCN 기업들이 수익성을 강화하고 지속 가능한 비즈니스 모델을 구축하기 위한 구체적인 전략적 방향을 제시함으로써 기업의 장기적인 성장과 발전을 도모하는 데 있어 중요한 참고 자료가 될 것이다. 본 연구에서 제시한 내용을 바탕으로, 향후 크리에이터 산업과 MCN 기업의 발전을 위한 후속 연구가 이어져야 할 것이다. 본 연구가 MCN 기업과 크리에이터 간의 협력 모델을 더욱 발전시키고 산업의 혁신을 촉진하는 데 이바지할 수 있기를 바란다.

참고문헌

[1] 문찬주. “초·중등 진로교육 현황조사(2023).” 교육부, 한국직업능력연구원, 2023.
 [2] 지봉철. “CJ E&M, 1인 콘텐츠 제작자 지원 나서.” 뉴스1, 2013.6.27.
 [3] 김하늘, 이해수, 박만수. “한국 MCN(Multi Channel Network) 기업들에 대한 연구, 사이버커뮤니케이션학보. 2018, 35(4):103-143.

문화적 푸시 및 풀 요인에 따른 여행 동기가 목적지 선택 만족도에 미치는 영향

강지효, 허지연, 차경진
한양대학교 경영대학 비즈니스인포매틱스학과
e-mail : wlgy2210@naver.com, jinny6290@gmail.com, kjcha@hanyang.ac.kr

Travel Motivation Based on Cultural Push and Pull Factors Impact on Destination Choices' Satisfaction

Ji hyo Kang, Ji yeon Hur, Kyung Jin Cha
Department of Business Informatics
School of Business, Hanyang University

Abstract

This research investigates the influence of cultural push and pull factors on tourists' destination choices and their travel satisfaction. By redefining the push-pull framework through the cultural context including psychological factors, the study provides empirical insights into how various motivations interact to shape travel decisions. Data from TripAdvisor reviews were analyzed using topic modeling and sentiment analysis to distinguish key travel motivations and their impacts on satisfaction. Results indicate that both internal cultural push factors (Asian and Western travelers) and external destination pull factors (Hedonic and Utilitarian attributes) have significantly different influences on travel motivations and satisfaction. Also, the precise interaction between these two factors has been demonstrated. The study discusses the importance of customizing marketing strategies to align with tourist preferences based on the cultural backgrounds and destination's attraction types.

1. Introduction

The major challenge for tourism industry is to identify the characteristics of individual visitors and to effectively promote the characteristics of their destination to them. (Goffi et al., 2018). To stimulate the desire about travelling specific destinations, travel motivation, is composed of a combination of intrinsic psychological push factors and extrinsic pull factors emerged from outside, is complex and multidimensional (Seebaluck et al., 2015). The push and pull framework helps explain how cultural background affects travel decisions and satisfaction (Valduga et al., 2020). However, previous studies projected cultural background as a factor influencing push and pull factors, but it should be regarded as the push factors itself, since social-psychological motives (push factors) stem from an individual's internal background (Iso-Ahola, 1983). In addition, it is important to consider the interaction between cultural push and pull factors, though research on this interaction has been limited. This study aims to reconstruct the new concept of push and pull motivations by finding out the individual effects and interaction effects of cultural push factors and pull factors of destination attributes on travel motivation, contributing to a better understanding of travel satisfaction.

2. Related works

2.1 Push and Pull Travel Motivations related to culture background

Push and pull motives have been proved to be pertinent and effective frameworks for understanding the reasons behind tourists' travel intentions and identifying their behaviors (Ayoub & Mohamed, 2024). Push factors are internal and psychological motivations that drive an individual to travel (Dale & Ritchie, 2020). Pull factors are external factors and refer to the needs and desires that can be met by people in the places they travel to (Dann, 1981). In most previous studies, the cultural background of travelers effects push or pull motivations as an influencing factor rather than defines as motivations itself. However, the traveler's psychology motives can vary based on cultural background since culture shapes common behavior patterns, beliefs, and norms, which significantly influence human thought and behavior (Berry, Poortinga & Pandey, 1997). Therefore, this study suggests that the push factor is redefined as a cultural background itself.

2.1.1 Cultural Push Factors: Asian or Western

Cultural background significantly influences travel motivations, with distinct differences observed between Asian and Western travelers. For instance, Asians often prefer group travel and familiar destinations, while Westerners lean towards individualistic and adventurous experiences (Hofstede, 1991).

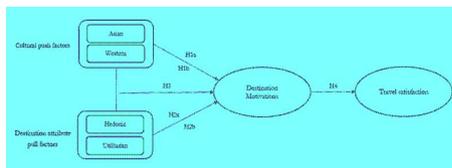
This study categorizes travelers based on these cultural dimensions to understand how these cultural push factors (Asian vs. Western) influence destination choices.

2.1.2 Destination Attribute Pull factors: Utilitarian or Hedonic

Destination attributes by attraction types are categorized into utilitarian and hedonic factors (Douglas et al., 2024). Utilitarian destinations offer practical, functional value, such as educational or business travel, while hedonic destinations provide pleasure and enjoyment, like beach resorts or entertainment hubs. These pull factors (Utilitarian vs. Hedonic) significantly influence travelers' destination choices and are shaped by cultural background (Byun & Jang, 2015).

2.2 Travel Satisfaction

Travel satisfaction is influenced by both push and pull factors, where understanding the individual effects of these motivations helps in enhancing overall tourist satisfaction (Susanti, 2021). Previous studies indicate that these factors can either positively or negatively impact satisfaction, making it essential to analyze them separately (Yoon & Uysal, 2005).



(Figure 1) The Proposed Model

3. Methods

In order to obtain destination motivations based push and pull affecting travel satisfaction from online reviews and to identify the impact of individual factors on the satisfaction, we propose a mixed approach that incorporates topic modeling (LDA) and extreme gradient boosting (XGBoost). In the first stage, it is possible to compare the distribution of motivations extracted through topic modeling by each cultural push factor and destination attribute pull factor. In the second stage, the impact of each motivational feature extracted in the previous stage on travel satisfaction can be quantified with strong sentiment score (1,5).

Acknowledgement

This study was supported that how push and pull factors influence travel destination choices and satisfaction, emphasizing the role of cultural background as a push factor. By analyzing tourist sentiments, it reveals that the importance of destination motivations between each four group does not always align with satisfaction levels. The study confirms both individual and interaction effects of cultural push and destination pull factors, providing valuable insights for tailoring marketing strategies to better meet the preferences of different groups's motivations.

References

- [1] Ayoub, D., & Mohamed, D. N. H. S., "The impact of push-pull motives on internal tourists' visit and revisit intentions to Egyptian domestic destinations: the mediating role of country image.", *Humanities and Social Sciences Communications*, 11(1), pp. 1-13, 2024.
- [2] Berry, John W., Ype H. Poortinga, and Janak Pandey, eds., "Handbook of Cross-cultural Psychology: Basic processes and human development", vol. 2, 1997.
- [3] Byun, J., & Jang, S. S., "Effective destination advertising: matching effect between advertising language and destination type." *Tourism Management* 50: pp. 31 - 40, 2015.
- [4] Dale, N. F., & Ritchie, B. W., "Understanding travel behavior: A study of school excursion motivations, constraints and behavior." *Journal of Hospitality and Tourism Management*, 43, pp. 11-22, 2020.
- [5] Dann, G. M., "Tourist motivation an appraisal.", *Annals of tourism research*, 8(2), pp. 187-219, 1981.
- [6] Douglas, A., Hoogendoorn, G., & Richards, G., "Activities as the critical link between motivation and destination choice in cultural tourism.", *Journal of Hospitality and Tourism Insights*, 7(1), pp. 249-271, 2024.
- [7] Goffi, G., Masiero, L., & Pencarelli, T., "Rethinking sustainability in the tour-operating industry: Worldwide survey of current attitudes and behaviors.", *Journal of cleaner production*, 183, pp. 172-182, 2018.
- [8] Hofstede G., "Cultures and Organizations: Software of the Mind.", McGraw-Hill: London., 1991.
- [9] Iso-Ahola, S. E., "Towards a social psychology of recreational travel.", *Leisure studies*, 2(1), pp. 45-56, 1983.
- [10] Seebaluck, N. V., Munhurrun, P. R., Naidoo, P., & Rughoonauth, P., "An analysis of the push and pull motives for choosing Mauritius as "the" wedding destination.", *Procedia-Social and Behavioral Sciences*, 175, pp. 201-209, 2015.
- [11] Susanti, C. E., "The effect of push motivational factors and pull motivational factors on tourist loyalty through satisfaction on halal tourists in east java.", *The Effect of Push Motivational Factors and Pull Motivational Factors on Tourist Loyalty Through Satisfaction on Halal Tourists in East Java*, 17, pp. 510-519, 2021.
- [12] Valduga, M. C., Breda, Z., & Costa, C. M., "Perceptions of blended destination image: the case of Rio de Janeiro and Brazil.", *Journal of Hospitality and Tourism Insights*, 3(2), pp. 75 - 93, 2020.
- [13] Yoon, Y., & Uysal, M., "An examination of the effects of motivation and satisfaction on destination loyalty: a structural model.", *Tourism management*, 26(1), pp. 45-56, 2005.

모바일 쇼핑몰의 환기성 높은 이미지가 원격현존감에 미치는 영향: VVIQ의 조절효과를 중심으로

김 협^{1*}, 권혁준², 차욱재³, 박지섭^{4†}

¹호서대학교 경영학부, ²순천향대학교 경제금융학과, ³성균관대학교
전자전기컴퓨터공학과, ⁴한국기술교육대학교 융합학과

e-mail : hyeob.kim@hoseo.edu, gloryever@gmail.com, wjcha2008@skku.edu,
jpark@koreatech.ac.kr

The Impact of Highly Evocative Images in Mobile Shopping Malls on Telepresence: Focusing on the Moderating Effect of VVIQ

Hyeob Kim^{1*}, Hyuk-Jun Kwon², Wook Jae Cha³, Jiseob Park^{4†}

¹Division of Business Administration, Hoseo University

²Department of Economics and Finance, Soonchunhyang University

³Department of Electrical and Computer Engineering, Sungkyunkwan University

⁴Department of Future Technology, Korea University of Technology and Education

요 약

본 연구는 모바일 쇼핑 환경에서 Virtualgraph 기술을 활용한 환기성 높은 이미지가 사용자의 원격현존감에 미치는 영향과 개인의 시각적 상상력 선명도(VVIQ)의 조절 효과를 검증하였다. Virtualgraph는 가상과 실제의 경계를 넘나드는 새로운 콘텐츠 제작 방식으로, VR 환경에서 사용자에게 높은 몰입감을 제공할 수 있다. 실험 설계 방법을 통해 Virtualgraph 이미지와 일반 정지 이미지의 효과를 비교하였으며, 5가지 과일 이미지를 자극물로 사용하였다. 104명의 참가자를 대상으로 실험을 진행한 결과, Virtualgraph를 활용한 환기성 높은 이미지가 일반 이미지보다 더 높은 원격현존감을 유발하는 것으로 나타났다. 또한, 개인의 VVIQ 수준이 환기성과 원격현존감 사이의 관계를 조절하는 것으로 확인되었다. 이러한 결과는 모바일 쇼핑 플랫폼 설계와 마케팅 전략 수립에 중요한 학문적, 실무적 시사점을 제공할 수 있을 것으로 사료된다.

1. 서 론

디지털 기술의 급속한 발전과 함께 모바일 쇼핑은 현대 소비자들의 일상에서 중요한 위치를 차지하게 되었다. 스마트폰의 보편화로 인해 시간과 장소에 구애받지 않고 쇼핑이 가능해졌지만, 실제 제품을 직접 경험할 수 없다는 한계점이 여전히 존재한다. 이러한 제약을 극복하고 사용자 경험을 향상시키기 위해 다양한 기술적 접근이 시도되고 있으며, 그중 하나가 Virtualgraph 기술이다.

Virtualgraph는 가상과 실제의 경계를 넘나드는 새로운 형태의 콘텐츠 제작 방식으로, VR 환경에서 사용자에게 높은 몰입감을 제공하는 것을 목표로 한다. 이 기술은 환기성(Evocativeness)을 높임으로써 사용자의 심상에 강력한 영향을 미치는 것으로 알려져 있다. 환기성은 특정 대상에 대한 다감각적 경험이 개인의 심상에 미치는 영향의

정도를 의미하며, 이는 심상이론을 기반으로 한다[1].

환기성은 수행 향상, 스트레스 감소, 우울 증상 완화 등 다양한 심리적, 정서적 효과를 포함하고 있다[2]. 특히, 선행 연구에서는 Virtualgraph가 사진과 동영상보다 환기성, 태도, 원격현존감, 기억력 부분에서 더 높은 효과를 보이는 것으로 나타났다[1].

한편, 개인의 시각적 상상력 선명도(Visual Imagery Vividness Questionnaire, VVIQ)는 Marks(1973)에 의해 개발된 이후 개인의 심상 능력을 측정하는 도구로 널리 사용되어 왔다. VVIQ는 인지 과정, 기억력, 창의성 등 다양한 심리적 구성 개념과의 관계를 탐구하는 데 활용되어 왔다[3]. 그러나 VVIQ와 환기성, 그리고 이들이 모바일 쇼핑 경험에 미치는 영향에 관한 연구는 아직 충분이 이루어지지 않았다.

본 연구는 모바일 쇼핑 환경에서 Virtualgraph 기술을 활용한 환기성 높은 이미지가 사용자의 원격현존감에 미치는 영향을 탐구하고, 이 과정에서 개인의 VVIQ가 어떤

* 제1저자(first author)

† 교신저자(corresponding author)

조절 효과가 있는지 검증하는 것을 목적으로 한다.

2. 관련 연구

2.1 환기성과 Virtualgraph

환기성(Evocativeness)은 특정 자극이 개인의 심상(mental imagery)에 미치는 영향의 정도를 의미한다. 이 개념은 이중부호화 이론(Dual Coding Theory)을 기반으로 하며, 비언어적 정보 처리 체계와 밀접한 관련이 있다. 이중부호화 이론에 따르면, 인간의 정보 처리는 언어적 체계와 비언어적 체계(이미지)를 통해 이루어지며, 두 체계가 상호작용하여 정보의 이해와 기억을 촉진한다.

환기성이 높은 자극은 더 풍부한 심상을 유발하여 정보 처리와 기억을 강화한다. 환기성 높은 광고는 소비자의 브랜드 태도와 구매 의도에 긍정적인 영향을 미친다. 이는 환기성이 높은 자극이 소비자의 심상 처리를 활성화시켜, 제품에 대한 더 생생한 경험을 제공하기 때문이다.

Virtualgraph는 이러한 환기성의 개념을 가상현실(VR) 기술에 적용한 새로운 콘텐츠 제작 방식이다[1]. 'Virtual(가상의·실제의)'과 'Graph(그리다)'의 합성어인 Virtualgraph는 가상 환경에서 실제와 유사한 경험을 제공하는 것을 목표로 한다. 이 기술은 기존의 정적인 이미지나 단순한 동영상보다 더 풍부한 감각적 경험을 제공함으로써 사용자의 심상을 강화한다.

2.2 시각적 상상력 선명도(VVIQ)

시각적 상상력 선명도(Visual Imagery Vividness Questionnaire, VVIQ)는 개인의 심상 능력을 측정하는 도구로, Marks(1973)에 의해 개발되었다. VVIQ는 개인이 마음속으로 이미지를 얼마나 선명하고 생생하게 떠올릴 수 있는지를 평가하는 척도로, 심리학, 인지과학, 신경과학 등 다양한 분야에서 활용되고 있다.

마케팅 및 소비자 행동 연구에서도 VVIQ의 중요성이 강조되고 있다. VVIQ 점수가 높은 소비자들이 광고 메시지에 더 강하게 반응하며, 제품에 대한 태도와 구매 의도가 더 높아진다는 것을 발견했다. 이는 VVIQ가 높은 개인이 광고를 통해 제시된 정보를 바탕으로 더 생생한 제품 사용 시나리오를 상상할 수 있기 때문으로 해석된다.

본 연구의 맥락에서, VVIQ는 Virtualgraph를 통해 제공되는 환기성 높은 이미지의 효과를 조절할 수 있는 중요한 개인차 변인으로 고려된다. VVIQ가 높은 개인은 Virtualgraph가 제공하는 풍부한 시각적 정보를 바탕으로 더 생생한 심상을 형성할 수 있으며, 이는 결과적으로 더 높은 수준의 원격현존감으로 이어질 수 있다.

2.3 모바일 쇼핑과 원격현존감(Telepresence)

모바일 쇼핑은 스마트폰과 태블릿 등 모바일 기기를 통해 상품을 검색하고 구매하는 활동을 의미한다. 최근 몇 년간 모바일 기술의 발전과 함께 모바일 쇼핑 시장은 급속도로 성장하였다. 그러나 모바일 쇼핑은 실제 매장에서

의 쇼핑 경험과 달리 제품을 직접 보고 만질 수 없다는 한계가 있다. 이러한 한계를 극복하고 사용자 경험을 향상시키기 위해 원격현존감(Telepresence) 개념이 주목받고 있다.

원격현존감은 Minsky(1980)에 의해 처음 제안된 개념으로, 매개된 환경에서 실제 환경에 있는 것과 같은 경험을 하는 정도를 의미한다. Steuer(1995)는 이를 가상현실 맥락으로 확장하여, "매개된 환경에 실제로 존재하는 듯한 느낌"으로 정의하였다. 원격현존감은 생동감(vividness)과 상호작용성(interactivity)이라는 두 가지 주요 요소에 의해 결정된다고 본다[4].

3. 연구 방법

본 연구는 모바일 쇼핑 환경에서 Virtualgraph 기술을 활용한 환기성 높은 이미지가 사용자의 원격현존감에 미치는 영향을 탐구하고, 이 과정에서 개인의 VVIQ(Visual Imagery Vividness Questionnaire)가 어떤 조절 효과를 갖는지 검증하기 위해 설계되었다. 이를 위해 실험 설계 방법을 채택하였으며, 정량적 데이터 분석을 중심으로 연구를 진행하였다.

4. 결론 및 시사점

본 연구는 Virtualgraph 기술을 활용한 환기성 높은 이미지가 모바일 쇼핑 환경에서 사용자의 원격현존감을 증진시키며, 개인의 VVIQ 수준이 이 효과를 조절함을 실증하였다. 이러한 결과는 모바일 쇼핑 플랫폼 설계에 있어 고급 이미지 기술의 중요성을 강조하며, 사용자의 시각적 상상력 특성을 고려한 맞춤형 마케팅 전략 수립의 필요성을 시사한다. 향후 다양한 제품군과 쇼핑 상황에서의 추가 연구가 필요할 것으로 사료된다.

참고 문헌

- [1] Park, J., Choi, J., Kim, H., & Kwon, H. (2019). The influence of media type and length of time delay on user attitude: Effects of product-focused virtual reality. *Computers in Human Behavior*, 101, 466-473.
- [2] Ardi, Z., Golland, Y., Shafir, R., Sheppes, G., & Levit-Binnun, N. (2021). The effects of mindfulness-based stress reduction on the association between autonomic interoceptive signals and emotion regulation selection. *Psychosomatic Medicine*, 83(8), 852-862.
- [3] McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery*.
- [4] Steuer, J. (1995). Defining virtual reality: Dimensions determining telepresence. In *Communication in the Age of Virtual Reality*. Lawrence Erlbaum and Associates.

방위산업기술 유출 현황 및 보안 대책 연구동향 분석

김성현¹, 임희영², 김자원*
대진대학교 IT융합대학 스마트융합보안학과
e-mail : {20231485¹, 20231509², jjawon*}@daejin.ac.kr

Analysis on Defense Industry Technology Leakage Status and Research Trends

Seonghyun Kim¹, Heeyoung Lim², Jawon Kim*
Dept. of Smart Convergence Security,
College of IT Convergence, Daejin University

요 약

한국의 방위산업기술 수준이 높아지고 있는 가운데, 해외 주요국에 수출액이 증가하는 등 방위산업의 중요성이 날로 커지고 있다. 기술의 수준이 높아짐에 따라 한국 방위산업기술의 유출 시도 또한 함께 증가하고 있으며, 최근에는 기술유출의 타겟이 방위산업이 되고 있다는 언론기사들도 쏟아지고 있는 실정이다. 이러한 위협에 대응하기 위한 방위산업 보호대책 또한 활발히 연구가 진행되고 있지만, 해외 수출 및 해외 생산 등과 관련해서 새로운 위협들이 등장하고 있다. 본 논문은 언론 보도를 기반으로 방위산업기술 유출 현황을 분석하고 유출 유형 대비 방위산업 기술보호 대책 연구동향의 추이를 살펴보고자 한다. 세부적인 연구 방법으로는 최근 5개년의 언론보도를 통해 발표된 방위산업기술 유출 시도의 유형을 분석하여 범주화하고, KCI 등제지를 기준으로 방위산업 기술보호 대책을 주제로 한 연구내용의 추이를 분석하고자 한다. 본 논문의 분석 결과는 향후 방위산업 기술보호 대책 수립의 방향성을 제시할 수 있는 기초자료로 활용될 수 있을 것으로 기대된다.

감사의 글

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음" (IITP-2024-RS-2024-00438056)

법제도 개선을 위한 텍스트 마이닝과 퍼지델파이 기반 장애인 분쟁유형 도출

황지선*, 윤태영**, 강주영***

*아주대학교 비즈니스애널리틱스학과, **아주대학교 법학전문대학원,

***아주대학교 경영인텔리전스 학과

e-mail : bibionidae@ajou.ac.kr, yty@ajou.ac.kr, jykang@ajou.ac.kr

Text Mining and Fuzzy Delphi Methods for Identifying Key Issues of Disabled for Legal Systems

Hwang Jisun*, Yoon Taeyoung**, Kang Juyong***

*Department of Business Analytics, Ajou University

**Law School, Ajou University

***Department of Business Intelligence, Ajou University

요 약

장애인복지법에 따르면, 장애인이란 신체적, 정신적 장애로 오랫동안 일상생활이나 사회생활에서 상당한 제약을 받는 사람을 말한다. 2023년 말 집계된 대한민국의 등록장애인 수는 264만 2,000명으로 총인구수 5175만명의 5.1%에 달한다. 등록되지 않은 장애인 수를 더한다면 그 수치는 더욱 증가할 것으로 예상된다. 우리 사회에서 장애인은 고용, 주거 등 생활에서 차별과 피해를 겪고 있으며 이에 대한 분석과 법제도 개선이 필요한 실정이다. 하지만 국내에서는 장애인의 법적 분쟁을 분석하여 유형화하고 법 제도의 개선 방향을 제안하는 연구는 부족한 상황이다. 따라서 본 연구에서는 텍스트 마이닝 기법을 활용하여 인터넷에서 언급되는 분쟁 유형을 분석하였다. 법률신문, 빅카인즈에서 신문 기사 데이터를 수집하여 토픽모델링을 통해 분쟁 유형을 도출하고 이를 유형화하였다. 이후 로톡과 네이버 지식인에서 수집한 상담 데이터를 통하여 분쟁 유형에 따른 법적 문제 사례를 도출하고 관련 법안의 수 요를 분석하였다. 본 연구는 토픽 선정 과정에서 법률 전문가들의 도움을 받아 주요 토픽 도출과 결과 해석을 진행하였다. 이때 퍼지델파이 기법을 활용하여 애매하거나 불확실할 수 있는 여러 전문가들의 의견들을 통합하여 토픽의 우선순위를 도출하였다. 본 논문은 장애인 관련 법 제도의 개선을 위한 분쟁 유형 도출 방법론을 제안하고자 하였으며 장애인뿐만 아니라 다양한 방면의 법제도 관련 분쟁 유형 또한 도출하여 법제도 개선 방향 수립에 기여할 수 있을 것으로 기대한다.

시계열 DB를 이용한 풍력발전 예측 머신러닝 시스템 개발

강명구¹, 정솔², 김민식², 정재윤^{1,2,*}

¹경희대학교 산업경영공학과, ²경희대학교 빅데이터응용학과
e-mail : {augrb, bohomi1995j, 2017100869. jjjung}@khu.ac.kr

Development of Machine Learning-based Wind Power Forecasting Systems Using Time-series DB

Myeong-Gu Kang¹, Sol Jeong², Minsik Kim², and Jae-Yoon Jung^{1,2,*}

¹Department of Industrial & Management Systems Engineering, Kyung Hee University

²Department of Big Data Analytics, Kyung Hee University

요 약

최근에 방대한 시계열 데이터를 분석하는 머신러닝 기법들이 개발되고 있다. 본 연구에서는 시계열 데이터베이스를 활용하여 풍력발전 예측을 위한 머신러닝 시스템을 개발하는 방법을 소개한다. 시계열 데이터베이스는 타임스탬프를 포함하는 레코드들을 저장하고 빠르게 검색 및 처리를 지원하는 시계열 데이터 전용 데이터베이스이다. 본 연구에서는 Machbase-NEO라는 시계열 데이터베이스를 사용하여 풍력발전 및 기상예보 데이터를 저장한다. 다음으로 다양한 머신러닝 기법을 적용하여 풍력 발전량을 예측하는 시스템을 구현하여, 시계열 데이터베이스의 활용 가능성에 대하여 검증하였다.

1. 서 론

시계열 데이터는 기상 정보, 사물인터넷 등 다양한 분야에서 시간의 흐름에 따라 수집되는 방대한 양의 데이터이다. 시계열 데이터를 효율적으로 저장하고 분석하는 시스템의 필요성이 증가하였다. 그러나 일반적인 관계형 데이터베이스(Relational DataBase Management System)는 대량의 시계열 데이터를 빠르게 저장하고 검색하는데 한계가 있다. 따라서, 시계열 형태의 데이터를 저장, 쿼리, 확장하기 위한 시계열 관리 시스템(Time Series Management Systems)을 개발 및 활용하고 있다[1].

관계형 데이터베이스 기반의 과거 시계열 관리 시스템은 대량의 시계열 데이터를 저장하고 질의하는데 최적화되지 않았다[2]. 그러나 최근에 개발된 시계열 데이터베이스(Time-Series DataBase)는 데이터와 데이터 처리 엔진과의 긴밀한 통합과 실시간 데이터 압축을 통해 데이터 저장 및 검색의 응답시간을 단축시킨다. 뿐만 아니라, 실시간 데이터 처리에서 데이터를 동시에 처리하고 시계열 데이터 특성에 따른 시계열 SQL 구문을 제공한다.

본 논문은 시계열 데이터베이스를 사용한 풍력 발전 예측을 위한 머신러닝 시스템을 개발하고, 기존 시스템과 비교하여 시계열 데이터베이스의 효과를 검증한다.

2. 풍력발전 예측 데이터

시계열 데이터베이스에는 기존에 확보한 기상 예보 데이터파일과 제주도 환경, 성산 두 지역의 풍력발전소의 발전 이력 데이터 파일[3]을 저장하였다. 두 데이터는 2014년부터 2017년까지 4년간 수집된 시계열 데이터이다. 총 63만여 건의 행에 총 390만여 건의 값을 포함한다.

기상 예보 데이터는 두 지역에 3시간 단위로 수집된 기온, 습도, 풍속, 풍향을 포함한 총 10개의 기상 예보에 대한 정보로 구성된다. 풍력발전 이력 데이터는 환경 지역의 두 발전소와 성산 지역의 발전소 내의 터빈의 발전량을 포함한다. 두 환경 발전소에는 각각 터빈이 3개, 5개, 성산 발전소는 10개의 터빈이 있다. 데이터파일에는 각 터빈의 풍력발전량이 10분마다 kW 단위로 수집되었다.

3. 시스템 아키텍처

시계열 데이터베이스를 사용한 머신러닝 기반 풍력발전 예측 시스템의 아키텍처는 그림 1과 같이 시계열 데이터베이스, 전처리 모듈, 학습 모듈, 추론 모듈로 구성된다. 시계열 데이터베이스에는 시계열 정보와 데이터 열과 열에 대한 값에 대해서 name, time, value로 구성된 트랙잭션 형태의 테이블을 저장한다. 이때 기존 데이터파일의 시계열 정보는 나노 초 단위의 에포크 시간으로 변경한다.

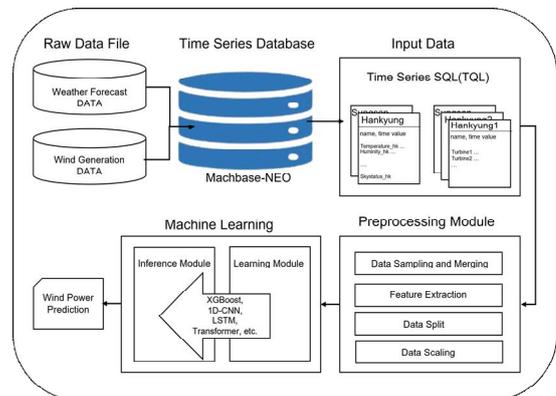


그림 1. 제안된 시스템 아키텍처

저장한 데이터를 불러오는 과정은 시계열 데이터베이스인 Machbase-Neo의 시계열 특성 SQL 구문인 TQL를 활용한다. TQL은 시계열 특성을 반영한 검색기반 쿼리를 통해 원하는 시간대와 원하는 이름의 값들을 데이터프레임 형태로 빠르게 불러온다. 또한, TQL 구문을 통해 서로 다른 시계열 정보가 동일한 시간대에 있다면, 값들을 다양한 형태로 수집 및 집계하여 가져올 수 있다.

데이터 전처리 모듈에서는 TQL 구문 및 파이썬 코드를 통해 데이터프레임을 생성한다. 이때, 데이터를 시간 영역을 집계하여 불러오거나, 두 데이터를 불러와 동일한 시간 값에 맞게 병합한다. 그 후 특성 생성, 정규화, 윈도우 슬라이싱 등의 전처리를 추가적으로 진행하여 모델 학습과 추론을 위한 데이터프레임 형태로 전처리한다.

학습 모듈과 추론 모듈은 XGBoost, Random Forest, 1D-CNN, LSTM 등 일반적인 머신러닝에서 딥러닝까지 다양한 머신러닝 모델을 활용한다. 시계열 예측 모형을 위해 다양한 특성 값으로 모델학습을 진행한다. 이후 추론 모듈로부터 테스트 데이터를 입력하여 원하는 시점의 풍력발전량에 대한 예측 값을 도출한다.

4. 실험 결과

학습 및 추론에는 기상 예보 정보와 시계열의 주기성, 과거의 발전정보 총 16개의 특성을 활용했으며, 목표변수는 각 시점의 터빈당 발전량이다. 기상예보 데이터 파일은 누락 값이 많은 열을 제외한 나머지 기온, 습도, 풍속, 강수확률을 사용하였다. 또한, 풍향과 풍속에 관련된 새로운 특성 3개를 추가하였다. 시계열의 주기성을 반영하는 코사인 변환 값으로 5개의 특성을 사용했으며, 추가적으로 계절과 공휴일을 구분하는 특성을 추가했다. 마지막으로 실질적으로 발전에 사용된 발전터빈의 개수, 터빈당 발전량을 최종적인 데이터프레임으로 사용하였다.

머신러닝 모델로 LSTM을 사용했으며, 과거 30일의 정보를 통해 한 시점(3시간) 후의 터빈당 발전량을 예측하는 One-Step 모형과, 여덟 시점(다음 날 하루) 후의 발전량을 예측하는 Multi Step 모형을 사용했다.

실험 결과 테스트데이터에 대한 One-Step 모형의 결정계수(R-squared)는 환경1, 환경2, 성산 발전소에서 각각 93.5%, 93.7%, 93.2%을 보였고, Multi-Step 모형은 각각 88.4%, 89.3%, 90%의 결정계수 값을 보여주었다.

시계열 데이터베이스 사용유무와 시스템 비용을 보기 위해 Machbase-NEO와 로컬 환경에서 두 데이터파일 불러와 합병하는 동안의 시스템 메모리 사용량과 소요된 시간을 측정하였다. 측정을 위해 파이썬의 time, psutil 라이브러리를 사용했다.

그 결과 오프라인 데이터 처리 방식의 경우 약 425MB의 시스템 메모리가 사용되었으며 60여 초 가량의 시스템 시간이 소요되었다. 반면, 시계열 데이터베이스를 사용한 경우 약 66MB의 메모리가 사용되었으며 7초 내외의 시간이 소요되었다. 원래 데이터파일을 Machbase-NEO에 올

릴 때 소비된 시스템 비용까지 합산하면 약 305MB의 시스템 메모리를 사용하고 12초 내외의 시간이 소요된다.

5. 결론 및 추후연구

본 연구에서는 머신러닝 기반의 풍력발전 예측 시스템을 구축하기 위하여 시계열 DB를 활용하는 방안을 제시하고, 활용 가능성을 분석하였다.

시계열 데이터베이스를 사용하는 경우 머신러닝 시스템의 데이터 처리 과정에서 시스템 비용의 유의미한 감소를 보였다. 특히 시계열 데이터베이스와 전처리 모듈 과정의 소요 시간에 큰 영향을 미쳤다. 이러한 이유는 Machbase-Neo에 데이터를 올릴 때 자동적으로 누락된 값들은 제외하기 때문이다. 또한, 시계열 특성 SQL 구문인 TQL의 쿼리를 통해 원하는 형태의 데이터프레임을 빠르게 전처리 모듈에 불러올 수 있었다. 추가적으로 웹브라우저 UI를 통해 데이터 테이블 내의 원하는 항목들을 그 래프와 같은 다양한 형태로 가시화할 수 있다.

다만, TQL 구문과 시계열 쿼리의 사용에 익숙해져야 하는 점과 쿼리의 한계로 인해 전처리 모듈의 모든 처리를 시계열 데이터베이스에서 수행하는 데 어려움이 있다.

앞으로 시스템 아키텍처의 모듈 상당부분을 대체할 수 있는 시계열 데이터베이스의 활용 방안과 그 효용성을 검증하는 것이 필요하다. 또한, 대규모 데이터와 실시간으로 수집되는 데이터들을 학습하고 추론할 수 있어야 한다. 따라서 대규모 데이터 기반의 새로운 풍력발전 예측을 위한 Pretrained Model 개발이 필요하다.

감사의 글

본 연구는 경희대학교 대학원 4단계 BK21 사업으로부터 교비 지원받은 연구임(GS-1-JX-NON-20241258).

참고문헌

- [1] S. K. Jensen, T. B. Pedersen, and C. Thomsen, "Time series management systems: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 11, pp. 2581-2600, Nov. 2017. doi: 10.1109/TKDE.2017.2740932.
- [2] Søren Kejser Jensen, Torben Bach Pedersen, Christian Thomsen. "Time Series Management Systems: A 2022 Survey." *Data Series Management and Analytics*, edited by Themis Palpanas, Kostas Zoumpatianos, Association for Computing Machinery, forthcoming.
- [3] M.-W. Baek, M. K. Sim, and J.-Y. Jung, "Wind power generation prediction based on weather forecast data using deep neural networks," *ICIC Express Letters, Part B: Applications*, vol. 11, no. 9, pp. 863-868, Sep. 2020.

LLM을 이용한 고장진단의 설명가능 AI 기술 개발

김민식¹, 최서현¹, 정재윤^{1,2*}

¹경희대학교 빅데이터응용학과, ²경희대학교 산업경영공학과
e-mail : {2017100869, choish91, jjjung}@khu.ac.kr

Developing Explainable AI for Fault Diagnosis Using Large Language Models

Minsik Kim¹, Seohyun Choi¹, and Jae-Yoon Jung^{1,2,*}

¹Department of Big Data Analytics, Kyung Hee University

²Department of Industrial & Management Systems Engineering, Kyung Hee University

요약

최근 산업 현장에서 불량 검출, 고장 진단, 공정 자동화 등 다양한 목적으로 머신러닝 기법을 활발히 도입하고 있습니다. 나아가, 산업 현장에서 사용되는 머신러닝 모델의 해석 가능성과 신뢰성을 높이기 위해 설명가능한 AI 기법이 개발되고 있습니다. 그러나 현재의 설명 가능한 AI 기법들은 여전히 기술적으로 복잡하며, 현장 담당자가 이해하고 활용하기에는 어렵습니다. 본 연구에서는 고장진단을 위한 머신러닝 모델을 설명하는 데 거대 언어 모델(Large Language Model: LLM)을 활용하는 시스템을 소개합니다. 이를 위해 OpenAI에서 제공하는 ChatGPT API를 사용합니다. ChatGPT와 같은 LLM을 통해 실무 담당자에게 머신러닝 모델의 작동 방식에 대한 질문과 답변을 제공할 수 있습니다. 제안된 시스템을 개발하기 위해, 먼저 진동 데이터를 활용해 고장 진단을 수행하는 머신러닝 모델을 학습시키고, OpenAI의 LLM인 "gpt-4o-mini-2024-07-18"을 사용하여 학습된 모델을 설명하는 기능을 구현하였습니다. 이 LLM 기반의 설명가능한 AI 시스템은 현장 담당자가 자연어로 산업 현장의 머신러닝을 이해하고 활용하는 데 도움을 줄 것으로 기대합니다.

감사의 글

본 연구는 산업통상자원부 재원으로 한국산업기술기획평가원[전자부품산업기술개발사업, No. RS-2024-00442354]과 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No. RS-2022-00155911, 인공지능융합혁신인재양성(경희대학교))의 지원을 받아 수행되었음.

자산관리셸 표준을 이용한 자율이동로봇 모니터링

신민중¹, 한영석¹, 정재윤^{1,2*}

¹경희대학교 빅데이터응용학과, ²경희대학교 산업경영공학과
e-mail : {shinmj, ysjhan, jyjung}@khu.ac.kr

Monitoring Autonomous Mobile Robots Using Asset Administration Shell

Min-Jong Shin¹, Young-Suk Han¹, and Jae-Yoon Jung^{1,2*}

¹Department of Big Data Analytics, Kyung Hee University

²Department of Industrial & Management Systems Engineering, Kyung Hee University

요 약

최근 자율이동로봇(Autonomous Mobile Robot: AMR) 모니터링을 위한 기술들이 발전하고 있다. 본 연구에서는 자산관리셸(Asset Administration Shell: AAS) 표준을 활용하여 AMR의 하드웨어 및 소프트웨어 정보를 체계적으로 관리하는 시스템을 소개한다. AAS는 물리적 및 디지털 자산을 통합 관리할 수 있는 표준을 제공하며, OPC-UA 프로토콜을 통해 실시간으로 AMR의 상태를 모니터링하고 데이터를 교환할 수 있게 한다. 본 연구에서는 자산의 상태를 JSON 형식으로 저장하고 송수신하며 이를 통해 유연한 데이터 관리와 실시간 모니터링이 가능한 시스템을 구축하였다. 본 연구는 AMR 모니터링을 위한 AAS 표준화 방법 및 활용에 도움이 될 것으로 기대한다.

1. 서 론

자율이동로봇(Autonomous Mobile Robot: AMR)은 주어진 환경에서 사람의 개입 없이 자율적으로 이동하며 작업을 수행할 수 있는 로봇을 의미한다. 최근 산업 현장에서 작업의 유연성과 효율성을 향상시키기 위하여 AMR 도입을 많이 검토하고 있다[1]. AMR 모니터링은 로봇의 위치, 경로, 상태, 작업 수행 여부 등을 실시간으로 관제하는 것이다. 이를 통해 잠재적 오류를 사전에 감지하고, 운용 효율을 극대화할 수 있다.

AMR의 운영을 효과적으로 관리하고 모니터링하기 위해서는 일관된 데이터 모델과 정보 구조가 필요하다. 자산관리셸(Asset Administration Shell: AAS) 표준은 독일 인터스트리 4.0에서 제안된 표준으로, 사이버-물리 시스템(Cyber-Physical Systems)의 디지털 표현을 통해 자산 정보를 관리하고 교환할 수 있는 구조를 제공한다[2]. 본 연구에서는 AAS 표준을 적용하여 AMR의 정보 모델을 설계하고, 이를 통해 AMR의 모니터링 및 관리 효율성을 높이는 방안을 제시하였다.

본 연구에서는 basyx 파이썬 라이브러리를 사용하여 AMR의 하드웨어 및 소프트웨어 요소들을 JSON 형식으로 표준화하였다. 또한, OPC-UA 통신으로 AMR의 현재 상태를 실시간으로 교환할 수 있는 시스템을 구축하였다.

2. AMR을 위한 AAS 표준 설계

2.1. AMR 정보 모델

AMR의 정보 모델은 로봇의 다양한 하드웨어 및 소

프트웨어 속성을 저장한다. AMR의 정보 모델은 AAS 표준의 서브모델들로 세분화하여 표현되었하였다. 예를 들어, CPU, GPU, 메모리 등의 하드웨어 사양은 Main_Module 서브모델로 구분하고, 거리 범위(distance_range) 및 각도 범위(angular_range)와 같은 라이더 센서 데이터는 별도의 서브모델로 표현하였다. 이러한 정보 모델은 데이터 통합 및 상호운용성을 고려하여 표준에 따라 설계하였다.

2.2. AMR 정보의 AAS 표현 설계

AAS를 사용하여 AMR의 물리적 및 디지털 자산을 통합 관리하였다. AAS 표준을 통해 자산의 속성(Property), 상태(State), 이력(History) 등을 메타데이터로 정의하고 이를 디지털 트윈으로 표현하였다. AMR의 좌표 정보는 실시간으로 업데이트되며, 하드웨어 및 소프트웨어 정보는 각각 서브모델로 세분화된다.

본 연구에서는 파이썬으로 구현된 basyx 라이브러리를 활용하여 AMR의 주요 하드웨어와 소프트웨어를 표현하는 AAS를 생성하고, 이를 JSON 및 XML 형식으로 직렬화하였다. Main_Module 서브모델은 CPU, GPU, 메모리 등의 하드웨어 사양을 정의하였고, Lidar_Sensor 서브모델은 센서 데이터 범위 및 해상도를 나타냈다. software_data 서브모델은 운영체제 및 소프트웨어 구성 요소를 포함하였다. 생성된 JSON 파일은 OPC-UA 프로토콜을 통해 서버로 전송되어 실시간 데이터 모니터링이 가능하게 한다. 이를 통해 AMR의 상태 변화나 작업 수행 정보를 실시간으로 관리하고, 다른 시스템과의 상호작용도 지원할 수 있었다.

3. 시스템 구현

AAS를 활용한 AMR 모니터링 시스템을 구현하기 위하여 몇 가지 중요한 기술적 구성 요소를 사용하였다. 이 시스템은 OPC-UA 프로토콜을 통해 AMR의 실시간 상태를 모니터링하고 데이터를 교환할 수 있도록 설계되었다. 서버와 클라이언트는 각각 opcua 라이브러리를 사용하여 구현하였으며, 이를 통해 AMR의 좌표 정보 및 기타 데이터를 실시간으로 송수신할 수 있게 하였다. 특히, JSON 형식으로 데이터를 송수신함으로써 AMR의 다양한 상태 정보를 보다 유연하게 관리할 수 있게 하였다.

OPC-UA 시스템에서는 노드(Node)라는 개념이 중요한 역할을 한다. 노드는 데이터를 저장하고 관리하는 기본 단위로, AMR의 위치, 배터리 상태와 같은 정보를 저장한다. 각 노드는 고유한 식별자를 가지며, 서버와 클라이언트는 이를 통해 데이터를 읽거나 수정할 수 있다[3].

3.1. OPC-UA 서버 구현

서버는 OPC-UA 서버로 설정되어 클라이언트로부터 데이터를 수신할 수 있도록 준비된다. opcua.Server 클래스를 사용하여 서버를 초기화하고, 서버의 엔드포인트를 설정하였다.

- 노드 설정: 서버의 루트 노드에서 Param 객체와 normal_operation 변수를 생성하였다. 이 변수에 클라이언트로부터 전달받은 x, y, z 좌표값을 저장한다. normal_operation.set_writable() 함수를 통해 클라이언트가 이 변수에 데이터를 쓸 수 있도록 설정하였다. 여기서 루트 노드는 다른 하위 노드들을 연결할 수 있는 기본 노드이다.
- JSON 데이터 관리: AMR의 상태를 유연하게 관리하기 위해 JSON 데이터를 저장할 수 있는 json_variable 노드를 추가하였다. 이 노드는 클라이언트로부터 JSON 형식의 데이터를 수신하며, 이를 파싱하여 다양한 상태 정보를 처리하고 저장한다.

3.2. OPC-UA 클라이언트 구현

OPC-UA 클라이언트로 설정되어 OPC-UA 서버에 연결된 후, AMR의 좌표 데이터를 주기적으로 서버로 전송한다. 클라이언트는 서버에서 제공하는 normal_operation 노드를 찾아, 해당 변수에 x, y, z 좌표 값을 지속적으로 업데이트한다.

- 연속적인 데이터 전송: OPC-UA 클라이언트는 서버에 연결된 상태에서 주기적으로 좌표 데이터를 송신한다. 서버는 이를 실시간으로 수신하고, 클라이언트에서 전송한 데이터를 처리한다. 클라이언트는 좌표 데이터뿐만 아니라 JSON 형식의 데이터도 서버로 전송할 수 있으며, 이를 통해 AMR의 상태를 더욱 상세하게 모니터링할 수 있다.

3.3. 통합 구조

OPC-UA 서버와 클라이언트 간의 통신을 통해 실시간 모니터링 시스템을 구축하였다. 이 시스템은 AMR의 상태(좌표 및 기타 센서 데이터)를 실시간으로 추적하며, 클라이언트로부터 전달된 데이터를 서버에서 수신 및 분석한다. 이러한 구조는 AMR의 효율적인 모니터링과 문제 발생 시 즉각적인 대응을 가능하게 한다.

특히, JSON 형식의 데이터를 수신하는 기능은 AMR의 다양한 상태 정보를 유연하게 처리할 수 있게 한다. 이는 시스템 확장 및 다양한 자산을 모니터링하는 데 유용한 기능으로, AAS 표준과 결합함으로써 AMR의 하드웨어 및 소프트웨어 정보를 통합적으로 관리하고, 실시간 상태를 모니터링할 수 있는 강력한 관리 도구로 발전시킬 수 있다.

4. 결론 및 추후연구

본 논문에서는 AAS 표준을 적용한 AMR 모니터링 시스템의 설계 및 구현 방안을 제시하였다. AAS를 사용하여 AMR의 체계적인 모니터링이 보다 체계적이고 효율적으로 이루어질 수 있음을 논의가능함을 보였다.

이 시스템은 AMR의 상태를 실시간으로 분석하고, 고장이나 이상을 조기에 탐지하여 효율적으로 대처할 수 있는 장점이 있다. 그러나 본 연구에서는 단일 AMR만을 대상으로 하였기 때문에, 다수의 AMR이나 다양한 자산을 동시에 관리하는 데는 한계가 있었다.

추후 연구에서는 AMR뿐만 아니라 다양한 자산을 연결하여 여러 자산을 동시에 모니터링하는 방법을 연구할 수 있다. 또한, 이상이 발생할 경우 서버에 접속해 원격으로 AMR을 제어할 수 있는 기능도 필요하다. 이를 통해 확장성 있는 자산 관리 및 모니터링 시스템을 개발할 수 있을 것으로 기대한다.

감사의 글

본 연구는 산업통상자원부 재원으로 한국산업기술기평과원[전자부품산업기술개발사업, No. RS-2024-00442354] 과 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No. RS-2022-00155911, 인공지능융합혁신인재양성(경희대학교))의 지원을 받아 수행되었음.

참고 문헌

- [1] U. Waseem, "Understanding AMR Robots: A Comprehensive Guide", WEVOLVER, October 2023.
- [2] J. R. Rahal, A. Schwarz, B. Sahelices, R. Weis, S. D. Antón, "The Asset Administration Shell as Enabler for Predictive Maintenance: A Review", X-MOL, November 2023.
- [3] Francesco Flammini, "Performance Analysis of OPC UA for Industrial Interoperability Towards Industry 4.0", MDPI Energies, August 2021.

수거 폐가전제품 정보관리 효율화를 위한 AI 기반 이미지 인식

이성훈, 김소현, 장태우
경기대학교 산업경영공학과

e-mail : scott40252@naver.com, sohyun2854@naver.com, keenbee@kgu.ac.kr

AI-Based Image Recognition to Streamline Information Management of Collected Waste Home Appliances

Seong-Hun Lee, So-Hyun Kim, Tai-Woo Chang
Department of Industrial and Management Engineering, Kyonggi University

초 록

본 연구는 폐가전제품을 수거하고 등록하는 시스템에서 품목 및 수량의 자동 확인 정확도 향상을 통해 증빙 검증 작업의 효율화를 목적으로 한다. 기존 수거 과정에서는 수거 매니저가 품목과 수량을 앱에 직접 기재하고 수거 사진을 촬영하여 증빙하였다. 현재 이러한 사진 증빙 검증 작업이 건별로 육안으로 확인하는 작업으로만 이루어져, 제한된 인력으로 대량의 증빙자료를 확인해야 하는 부담이 발생하고 있다. 이를 해결하기 위해, 본 연구에서는 객체 탐지 기술인 YOLO(You Only Look Once) 모델을 활용하여 수거 제품을 자동으로 인식하는 시스템을 개발하였고, 다양한 폐가전제품의 탐지 및 인식률을 평가하였다. YOLOv8 large 모델을 적용한 결과, mAP0.5-0.95 값이 0.892로 높은 성능을 얻을 수 있었다. 인공지능을 활용한 수거 폐가전제품 인식 시스템이 등록된 증빙 사진 확인 작업의 효율성을 크게 향상시킬 수 있음을 확인하였다. 또한, 본 시스템은 향후 더 많은 데이터 학습을 통해 성능이 지속적으로 개선할 예정이며, 전기전자제품 폐기 관리 프로세스의 자동화를 위한 중요한 기반을 제공할 것이다. 추가로, 수거 당시 등록한 사진의 특성상 탐지하고자 하는 객체가 다양하고 배경이 복잡한 상황으로 인해 객체의 정확한 탐지가 어려운 문제가 존재하고, 인식의 정확도를 높여 문제를 해결하기 위한 추가 학습 방안도 함께 제시한다.

생성형 AI 기반 법률 지원 챗봇: 금융, 재난 관리 및 의료 분야에서의 BERT 기반 데이터 처리 및 응용에 대한 연구

박태윤, 유현욱, 김무철*
중앙대학교 컴퓨터공학과
e-mail : {asdf5178, yu990410, kimm}@cau.ac.kr

Generative AI-Based Legal Support Chatbot: A Study on BERT-Driven Data Processing and Application in Finance, Disaster Management, and Healthcare

Taeyoon Park, Hyunwook Yu, Mucheol Kim*
Chung-Ang University

Abstract

This study proposes a customized legal Q&A chatbot to address legal issues faced by the elderly in an aging society. Utilizing large language models (LLM) and natural language processing (NLP) technologies, this chatbot enables the elderly to describe their legal issues in natural language without prior knowledge of legal terms or procedures. The proposed system includes natural language search and multi-turn search functions, providing consistent legal information with high accuracy across various scenarios. This aims to support the elderly in independently managing legal issues, thereby improving their quality of life.

1. Introduction

As society ages, the elderly face various legal issues such as retirement, long-term care, medical decisions, and wills[1], [2]. However, current legal services are largely inaccessible to the elderly, who are considered a digitally vulnerable group. This difficulty in accessing legal support often leaves them struggling to resolve legal matters on their own, which can negatively impact their quality of life.

The advancement of artificial intelligence technology can play a crucial role in addressing these challenges. Particularly, large language models (LLMs) have achieved groundbreaking results in the field of natural language processing (NLP)[3]. Trained on vast amounts of text data, LLMs have acquired human-level language comprehension and generation capabilities[4]. Leveraging these technologies, this study proposes a customized legal Q&A chatbot designed specifically for the elderly.

This chatbot allows users to describe their legal issues in natural language without prior knowledge of legal terms or procedures. The AI then analyzes these descriptions to provide appropriate legal information, offering practical assistance in various fields such as finance, disaster management, and healthcare.

2. Related works

LLMs have achieved significant breakthroughs in the field

of natural language processing, with GPT-3 demonstrating high accuracy in legal Q&A tasks [5], [6]. Additionally, various models have been proposed in the field of sentence embedding. For instance, SBERT (Sentence-BERT) modifies BERT into a Siamese network structure to efficiently compare semantic similarities between sentences [7], while SimCSE uses unsupervised learning techniques to learn sentence similarities, applying standard dropout to create two different embeddings as positive pairs [8]. These models effectively capture the meaning of complex texts such as legal documents.

3. Methodology

This study proposes a specialized legal Q&A chatbot service for the elderly. The core elements of the service are natural language search and multi-turn search systems.

3.1 Natural Language Search Service

To enhance search convenience, a chat-based service was designed. By fine-tuning LLMs based on legal data and user Q&A data, the system provides appropriate legal information even if users input questions without prior knowledge of legal terms or procedures.

3.2 Multi-Turn Search System

The multi-turn search service resolves uncertainties. By embedding words and sentences into high-dimensional vectors and placing semantically similar content close together, the system understands user questions and provides appropriate

answers. Additionally, a network embedding model for detecting multiple search results through several rounds of conversation allows the system to accurately grasp user needs and provide customized legal information. This study aims to build a system based on these methodologies and evaluate its performance through various experiments. The goal is to offer a system that enables the elderly to independently solve legal issues.

4. Experiment

The proposed specialized legal Q&A chatbot service for the elderly was implemented and tested

4.1 Q&A on Laws Regarding Local Tax Payment

In Table 1, when asked about personal local tax payment, the system provided information about local tax laws. When asked about corporate tax payment, it provided information about corporate tax laws. This demonstrates the system's ability to understand the context of various user questions and provide appropriate legal information

(Table 1) Legal Q&A Chat Service, "Tax"Keywords summary

number	speaker	content
1	Human	Payment of Local Taxes for Individuals
2	Model	Local Tax Act, Article 150 (Taxpayer) (...) Taxpayers of local education tax are as follows: (...) 5. Taxpayers of individual and business establishment shares of resident tax.
3	Model	Local Tax Act, Article 118-2 (Deferment of Payment) 1 (...) The head of a local government may allow deferment of the payment of property tax on a single household single residence (...) if a taxpayer meeting all the following requirements applies for deferment three days before the expiration of the payment deadline. (...) 4. The amount of property tax to be deferred on the residential property for the relevant year exceeds 1 million KRW.
4	Human	Payment of Taxes for Corporations
5	Model	Local Tax Act, Article 150 (Taxpayer) (...) The taxpayers of local education tax are as follows. (...) 5. Taxpayers of individual and business establishment shares of resident tax.
6	Model	Corporate Tax Act, Article 3 (Taxpayer) (...)
7	Model	Corporate Tax Act, Article 3 (Taxpayer) (...)



(Figure 1) Legal Q&A Chat Service, "Tax"Keywords Original

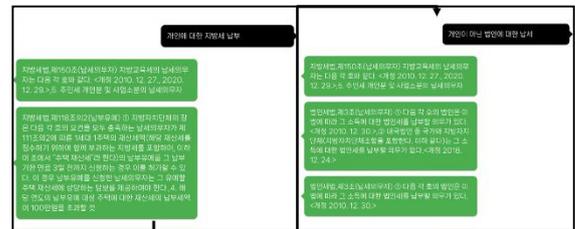
4.2 Q&A on Laws Regarding Hazardous Gases

In Table 2, when asked about laws regarding hazardous gases, the system retrieved regulations on fire safety management for public institutions and provided answers about the Special Act on Disaster Management for high-rise and

underground complex buildings. When subsequently asked about safety management laws for "very large buildings" instead of "high-rise," the system once again provided legal information on high-rise buildings, naturally continuing from the previous content

(Table 2) Legal Q&A Chat Service, "Hazardous Gas"Keywords summary

number	speaker	content
1	Human	Law on Harmful Gases
2	Model	Regulations on Fire Safety Management of Public Institutions, Article 5 (Appointment of Fire Safety Managers) 1. The head of an institution must appoint a qualified person in a supervisory position as a fire safety manager. However, public institutions that only install fire extinguishers or emergency alarm systems (...) may not need to appoint a fire safety manager.
3	Model	Special Act on Disaster Management of High-Rise and Underground Complex Buildings, Article 5 (Relationship with Other Laws) This law takes precedence over other laws concerning disaster and safety management of high-rise buildings, etc.
4	Human	Safety Management Law for Very Large Buildings
5	Model	Local Tax Act, Article 150 (Taxpayer) (...) The taxpayers of local education tax are as follows. (...) 5. Taxpayers of individual and business establishment shares of resident tax.
6	Model	Special Act on Disaster Management of High-Rise and Underground Complex Buildings, Article 9 (...)
7	Model	Special Act on Disaster Management of High-Rise and Underground Complex Buildings, Article 17 (...)



(Figure 2) Legal Q&A Chat Service, "Hazardous Gas"Keywords Original

5. Conclusion

This study proposes a specialized legal Q&A chatbot service to address the legal needs of the elderly. By leveraging advanced AI technologies such as large language models and graph neural networks, this service provides real-time, accurate legal advice through an intuitive chat interface. This allows the elderly to receive legal support without prior legal knowledge. The user-friendly interface and continuous feedback mechanisms help the elderly manage their legal issues independently. By addressing important legal topics such as inheritance and medical decisions, this service is expected to significantly improve the quality of life for the elderly, enhancing their independence and confidence in handling legal matters.

Acknowledgement

This Research was supported by the Korea Institute of Science and Technology Information(KISTI, P24023) and in part by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2024-RS-2024-00438056) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

References

[1] D. E. Bloom et al., "Macroeconomic implications of population ageing and selected policy responses", *Lancet*, vol. 385, no. 9968, pp. 649 - 657, Nov. 2014.

[2] M. Riswan and M. T. R. Farwin, "Challenges faced by elders: a critical evaluation," *KALAM - International Journal*, vol. 14, no. 2, 2021.

[3] J. D. M.-W. C. Kenton and L. K. Toutanova, 'Bert:

Pre-training of deep bidirectional transformers for language understanding', in *Proceedings of naacL-HLT*, 2019, vol. 1, p. 2.

[4] H. Naveed et al., 'A comprehensive overview of large language models', *arXiv preprint arXiv:2307.06435*, 2023.

[5] W. X. Zhao et al., 'A survey of large language models', *arXiv preprint arXiv:2303.18223*, 2023.

[6] H. Zhong, C. Xiao, C. Tu, T. Zhang, Z. Liu, and M. Sun, 'How does NLP benefit legal system: A summary of legal artificial intelligence', *arXiv preprint arXiv:2004.12158*, 2020.

[7] N. Reimers, 'Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks', *arXiv preprint arXiv:1908.10084*, 2019.

[8] T. Gao, X. Yao, and D. Chen, 'Simcse: Simple contrastive learning of sentence embeddings', *arXiv preprint arXiv:2104.08821*, 2021.

그래프 임베딩 개선을 위한 랜덤 워크의 내재된 편향성 완화 연구

류예중, 김무철*
중앙대학교 컴퓨터공학과
e-mail : {dpwhd621, kimm}@cau.ac.kr

Reducing Bias in Random Walks for Optimizing Graph Embeddings

Yejong Ryu, Mucheol Kim*
Department of Computer Science and Engineering, Chung-Ang University

요 약

랜덤 워크는 그래프 상에서 노드 간의 유사성을 학습하고, 그래프의 구조적 특성을 반영할 수 있는 그래프 임베딩에 분야에 효과적인 모델이다. 그러나 랜덤 워크 방법론들은 네트워크 토폴로지에서 중요한 다리 역할을 자주 하는 중심성이 높은 노드를 중심으로 학습하는 편향을 보이는 경향을 보인다. 이러한 편향은 학습 과정에 불균형적인 영향을 미쳐 그래프의 전체적인 임베딩을 왜곡할 수 있으며, 특히 연결이 드문 고립된 노드에 대해서는 부정확한 예측을 초래할 수 있다. PageRank는 그래프 내에서 노드의 역할과 연결 구조에 따라 중요도를 결정하는 알고리즘으로, 이를 통해 각 노드가 차지하는 상대적 중요도를 정밀하게 측정할 수 있다. 논문에서는 PageRank 알고리즘을 활용하여 노드의 중요도를 정량화하고 임베딩 과정에서의 불균형을 식별하는 새로운 접근 방식을 제안한다. 이러한 방법론을 통해 랜덤워크 기반의 그래프 임베딩에서 발생할 수 있는 편향을 완화하고, 모델의 표현적 공정성과 예측 정확성을 향상시킨다.

감사의 말

이 연구는 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0020632, 2024년 산업혁신인재성장지원사업).

합성 데이터셋 생성 및 이미지 분류를 이용한 도심 침수 탐지

강운제*, 정용훈**, 서상현***

중앙대학교 예술공학대학 컴퓨터예술학부*

중앙대학교 예술공학부***

중앙대학교 예술공학대학원 응용예술공학과**

e-mail : bluejay100@cau.ac.kr, dydgn2017@cau.ac.kr, sanghyun@cau.ac.kr

Urban Flood Detection Based On Synthetic Dataset Generation And Image Classification

YoonJe Kang*, Yonghoon Jung**, Sanghyun Seo*

The College of Art and Technology, Chung-Ang University*

Applied Art and Technology, Chung-Ang University**

요약

본 연구는 도시 침수 탐지를 위한 합성 데이터를 생성하는 파이프라인을 제안한다. 또한, 이미지 분류 모델을 활용하여 생성된 합성 데이터의 유용성을 측정하였다. 실제 및 합성 데이터를 혼합하여 사용한 결과, 정밀도 95.14%, 재현율 90.74%, 정확도 93.11%, F1-Score 92.89%의 성능을 보였다. 이는 실제 데이터만 사용한 결과보다 개선된 성능을 보여준다. 본 연구의 합성 데이터셋은 도심 침수 데이터 부족 문제를 해결할 것으로 기대한다.

1. 서론

지구 온난화로 인해 전 세계적으로 도심 침수가 빈번하게 발생한다. 이는 인류의 삶과 도시 인프라에 심각한 위협이 된다. 최근, 이를 예방하고 관리하기 위해 딥러닝 기반의 지도 학습 방법을 활용한 연구가 진행되고 있다. 그러나, 이러한 방법은 잘 구축된 실세계 기반의 데이터에 크게 의존하여 상당히 많은 경제적, 시간적 비용이 소요된다.

이를 해결하기 위해 우리는 3D 엔진을 사용하여 도심 침수 탐지를 위한 합성 데이터셋을 구축하는 파이프라인을 제안하며, 생성된 합성 데이터의 유용성을 평가하였다.

2. 데이터 구축 및 성능 평가 파이프라인

2.1 실제 데이터셋 수집



(그림 1) 실제 침수 데이터 샘플

실제 데이터셋은 차량 기반 침수 탐지에 관한 선행 연구에서 가져왔다. 해당 데이터셋은 차량의 침수된 비율에 따라 5단계로 나눈 라벨 정보를 포함한다[1].

그림 1은 실제 데이터셋에 포함된 데이터 중 일부이다. 실제 및 합성 데이터의 클래스 정보를 통일하기 위해 실제 데이

터셋의 클래스를 “침수”와 “비침수” 두 가지로 재분류하였다.

2.2 합성 데이터셋 구축

합성 데이터 생성 과정은 그림 2와 같다. 실제와 유사한 가상 도시 환경을 조성하기 위해 수집한 3D 에셋들을 통합하여 제작했다. 또한, 학습 데이터의 다양성을 높이고자 Nvidia Omniverse Isaac Sim에서 제공하는 랜덤화 기능을 활용하였다.

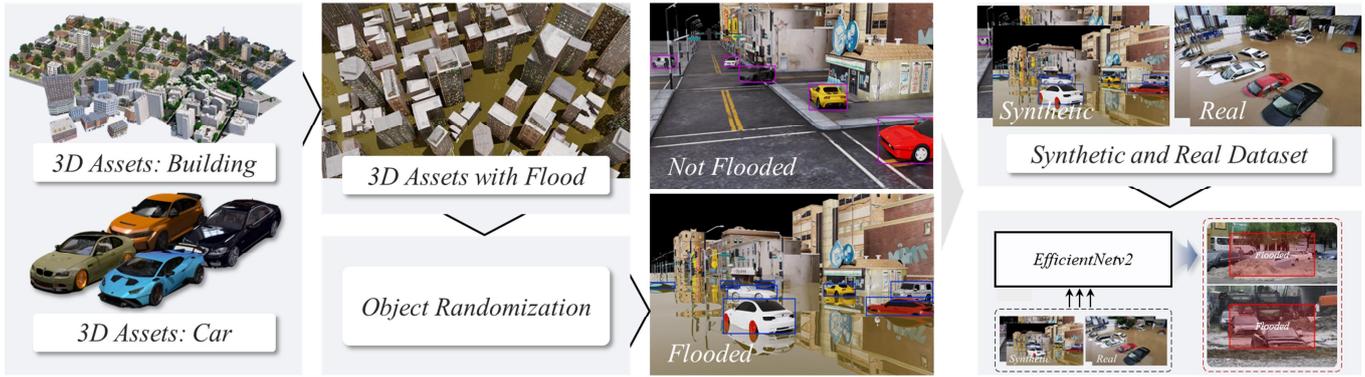
마지막으로 합성 데이터 생성 시 “침수”와 “비침수” 클래스로 분류하여 생성하였다. 구축된 합성 데이터셋은 표 1과 같이 침수 이미지 5,590장, 비침수 이미지 5,870장으로 총 11,460장의 이미지로 구성된다. 그림 3은 생성된 합성 데이터 일부를 보여준다.

(표 1) 합성 데이터셋 구축 결과

분류	이미지(장)
침수	5,590
비침수	5,870
합계	11,460

2.3 합성 데이터셋 유용성 평가

우리는 합성 데이터셋의 유용성을 평가하기 위해 이미지 분류 모델 EfficientNetv2을 사용하였다[2]. 실제 데이터와 합성 데이터를 독립적으로 모델 학습 및 테스트 과정을 진행하였다. 이는 합성 데이터의 유용성을 평가하는 것으로 두 모델은 동일한 실제 데이터로 테스트하였다.



(그림 2) 가상 침수 환경 구현 및 합성 데이터 생성 파이프라인

3. 실험 결과

EfficientNetv2 모델의 학습 파라미터는 다음과 같다. Batch size 8, 학습률 0.001, epoch 100, dropout rate 0.3, 그 외의 하이퍼파라미터는 기존 모델과 동일하다. 학습에 사용된 하드웨어 환경은 48 Core Intel Xeon(Cascadelake) 2.4GHz Processor와 Nvidia RTX A4500 8EA로 구성되었다.

실험 결과, 표 2과 같이 실제 데이터가 합성 데이터보다 성능이 좋았다. 합성 및 실제 데이터를 혼합한 경우, 모든 지표에서 성능 향상이 이루어졌다.

(표 2) 학습된 실제 데이터와 합성 데이터 모델의 성능 지표 비교

Method	Precision	Recall	Accuracy	F1-Score
Real	87.85%	87.03%	87.61%	87.44%
Synthetic	88.63%	36.11%	66.05%	51.31%
Synthetic and Real	95.14%	90.74%	93.11%	92.89%

향후 연구에서는 다양한 조건을 반영한 합성 데이터를 생성하여 품질을 개선하는 것을 목표로 할 것이다.

감사의 글

이 연구는 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0020632, 2022년 산업혁신인재성장지원사업)

참고문헌

[1] J. Wan, Y. Qin, Y. Shen, T. Yang, X. Yan, S. Zhang, G. Yang, F. Xue, and Q. J. Wang, "Automatic detection of urban flood level with YOLOv8 using flooded vehicle dataset," Journal of Hydrology, vol. 639, 2024.
 [2] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," Proc. Int'l Conf. Machine Learning (ICML), PMLR, 2021.



(그림 3) 합성 데이터 샘플

4. 결론

본 연구에서는 도심 침수 탐지를 위한 합성 데이터 생성 파이프라인을 제안하고 이미지 분류 모델을 사용해 구축된 합성 데이터셋의 유용성을 평가했다. 실험 결과, 합성 데이터와 실제 데이터의 혼합 사용으로 성능 향상이 이뤄졌다. 따라서 합성 데이터는 실제 데이터를 구하기 어려운 상황에서 좋은 대안이 될 수 있다. 이 연구는 도시 관리 및 재난 대응 시스템의 효율성을 향상시켜 궁극적으로 실제 침수 발생 시 신속한 대응과 피해 최소화에 기여할 수 있다.

메타버스 내의 디지털 어뷰징 행동 탐지를 위한 학습 데이터 세트 구축 연구

조용규*, 정용훈**, 서상현***

중앙대학교 예술공학대학 컴퓨터예술학부*

중앙대학교 예술공학부***

중앙대학교 예술공학대학원 응용예술공학과**

e-mail : justin0701@cau.ac.kr, dydgns2017@cau.ac.kr, sanghyun@cau.ac.kr

Research On Building Training Datasets For Detecting Abusing Behavior In the Metaverse

Yonggyu Cho*, Yonghoon Jung**, Sanghyun Seo*

The College of Art and Technology, Chung-Ang University*

Applied Art and Technology, Chung-Ang University**

요 약

실재감 있는 메타버스가 성범죄, 폭력 등 부적절한 행동의 배경이 된다. 본 연구에서는 이러한 행동을 탐지하기 위해 메타버스를 기반으로 한 어뷰징 학습 데이터 세트를 구축하였다. 이를 위해 메타버스에서 부적절한 행동들을 정의한 후, 프롬프트 기반 3D 행동 생성 확산 모델과 모션 캡처를 사용하여 3D 어뷰징 행동 데이터를 생성하였다. 이후, 생성된 행동 데이터를 메타버스 캐릭터로 리타겟팅 시켜 RGB 비디오 데이터를 추출했을 뿐만 아니라, 자동 라벨링 기법을 활용하여 메타버스 속 캐릭터의 어뷰징 행동을 인식 하기 위한 캐릭터의 바운딩 박스 및 2D 관절 좌표를 추출하여 학습 데이터 세트를 생성하였다. 이는 다양한 메타버스 상황과 어뷰징 행위 유형을 포괄하고, 이를 통해 어뷰징 탐지 모델을 효과적으로 학습시켜 메타버스의 건전한 환경 조성에 기여할 것으로 기대된다.

1. 서 론

메타버스는 빠르게 성장하며 현실과 유사한 사회적 상호작용을 제공하지만, 부정적인 행동 또한 증가하고 있다[1]. 이러한 문제를 해결하기 위해서는 딥러닝 기반 자동 탐지 시스템이 필요하나, 현재 어뷰징 행동 탐지를 위한 학습 데이터는 부족하며, 실제 데이터 사용에는 윤리적 문제가 따른다. 해당 문제를 해결하기 위해 본 연구는 메타버스 어뷰징 학습 데이터를 구축하기 위한 파이프라인을 제안한다. 프롬프트 기반 3D 행동 생성 확산 모델(MDM)[2]과 모션 캡처 기법을 활용해 3D 어뷰징 행동 데이터를 생성하고, 이를 게임엔진 Unity를 활용하여 캐릭터에 리타겟팅하여 학습 데이터 세트를 구축한다. 또한, 메타버스 내 어뷰징 행동 탐지를 위한 캐릭터의 바운딩 박스와 2D 관절 좌표 데이터를 자동으로 추출하여 활용한다[3].

본 연구는 메타버스 환경에서 어뷰징 탐지 학습 데이터를 효과적으로 구축하는 방법을 제시하고, 데이터 수집 문제를 개선한다.

2. 어뷰징 학습 데이터 세트 구축

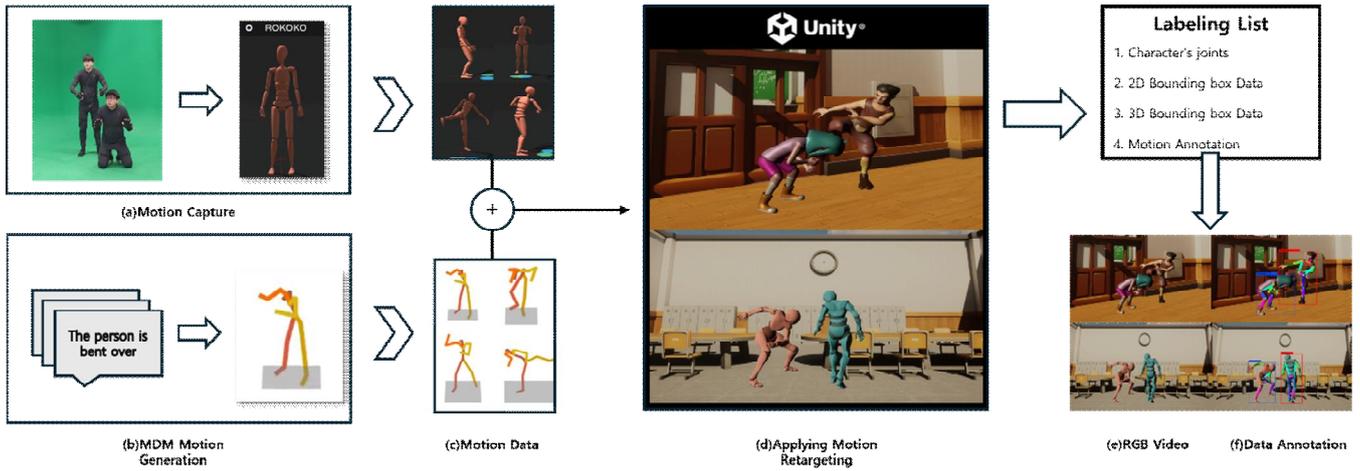
2.1 메타버스 내 어뷰징 행동 정의

메타버스의 자유로운 환경은 다양한 행동을 가능하게 하며, 본 연구는 디지털 어뷰징의 부적절한 행동에 초점을 맞추고 있다. 현재 국내외 법률 체계에서 메타버스에 대한 법적 정의가 부족하기 때문에 이러한 행위를 정의하고 분류

하는 것이 중요하다. 본 연구에서는 디지털 어뷰징 행위를 유사 성행위, 폭력 행위, 방해 행위의 세 가지 주요 범주로 분류하고, 가상의 이미지와 영상 자료를 제작하여 설문조사를 실시했다. 설문조사 결과, 메타버스 내 제재 필요성에 대해 과반수 이상의 이용자가 동의하는 것으로 나타났다. 이후, 설문조사 결과를 기반으로 디지털 어뷰징 행동을 세부적으로 정의하여 메타버스 내 어뷰징 행동을 구체화했다.

2.2 어뷰징 학습 데이터 세트 구축 파이프라인 설계

본 연구에서 제시한 메타버스 어뷰징 학습 데이터 세트 구축 파이프라인은 그림 1에 설명되어 있다. 현재 시중에는 다양한 3D 행동 데이터가 존재하지만 본 연구의 주요 초점인 어뷰징 행동을 반영하는 데이터는 부족하다. 이러한 데이터를 수동으로 생성하는 과정은 많은 시간과 비용이 소요되며, 실제 사람의 데이터를 사용할 경우 사생활 침해와 윤리적 문제가 발생할 수 있다. 이러한 문제를 해결하기 위해 우리는 사전에 정의된 어뷰징 행동을 시나리오별로 묶어 Rokoko에서 제공하는 모션 캡처 기법을 활용해 3D 어뷰징 행동 데이터를 생성하였다. 이는 그림 1-(a)에서 확인할 수 있다. 그림 1-(b)는 MDM을 사용하여 간단한 프롬프트를 기반으로 3D 어뷰징 행동 데이터를 생성해, 모션 캡처에서 부족한 시나리오별 어뷰징 행동 데이터를 보완하는 과정을 보인다. 이렇게 생성된 두 가지 3D 어뷰징 행동 데이터를 결합한 결과는 그림 1-(c)에서 확인할 수 있다.



(그림 1) 어뷰징 학습 데이터 세트 구축 파이프라인

생성된 3D 어뷰징 행동 데이터는 게임 엔진 Unity에서 각 캐릭터별로 리타겟팅되어, 메타버스 내 다양한 캐릭터의 행동을 반영하도록 했다. 3D 어뷰징 행동 데이터가 적용된 캐릭터들을 다양한 배경과 결합하여 사전에 정의된 어뷰징 행동 시나리오를 구축하고, 이를 렌더링해 RGB 비디오 학습 데이터 세트를 생성했다. 그 결과는 그림 1-(d)에서 확인할 수 있다.

마지막으로 Unity의 perception Package[4]를 활용하여 Top-Down 방식 행동 인식 연구를 위한 데이터를 추출했다. 이는 어뷰징 시나리오에서 각 캐릭터가 3D 어뷰징 행동을 취할 때 자동으로 캐릭터의 바운딩 박스 좌표 값과 2D 관절 좌표 값을 추출한다. 이를 활용하여 생성된 메타버스 내 어뷰징 행동 학습 데이터는 그림 2에서 확인할 수 있다.

3. 결론

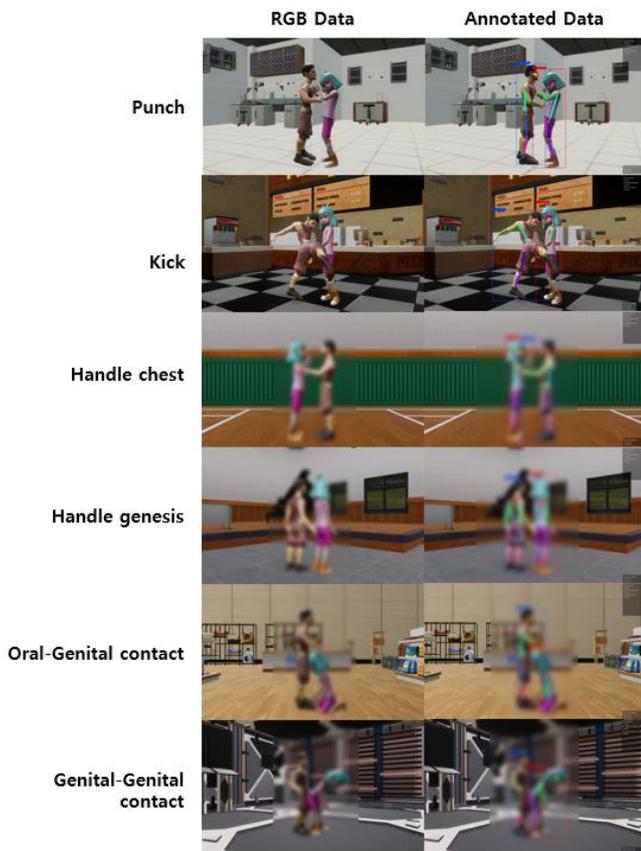
본 연구는 메타버스에서 어뷰징 행동을 탐지하기 위한 학습 데이터 구축 파이프라인을 제안한다. MDM과 모션 캡처 기법을 활용해 디지털 어뷰징 행동 데이터를 생성하고, 이를 Unity를 사용해 다양한 캐릭터에 리타겟팅하여 어뷰징 시나리오별 RGB 비디오와 다양한 라벨링 정보를 포함한 학습 데이터 세트를 추출했다. 이 연구는 메타버스 환경에서 어뷰징 행동 학습 데이터를 효과적으로 구축하는 방법을 제시하여, 합성 데이터 수집의 어려움을 개선하고 메타버스의 건전한 환경 조성에 기여할 것으로 기대된다.

감사의 글

이 연구는 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0020 632, 2022년 산업혁신인재성장지원사업)

참고문헌

[1] The Metaverse Is Coming And It's A Very Big Deal, Forbes, Jul. 5, 2020.
 [2] Guy Tevet, Sigal Raab, Brian Gordon, Yoni Shafir, Daniel Cohen-or, and Amit Haim Bermano, "Human Motion Diffusion Model," 2023, ArXiv abs/2209.14916.
 [3] Hua, Zhixin, Zheng Wang, Xingshi Xu, Xiangfeng Kong and Huaibo Song, "An effective PoseC3D model for typical action recognition of dairy cows based on skeleton features," Comput. Electron. Agric. 212: 108152, 2023.
 [4] Erfanian Ebadi, Salehe, Saurav Dhakad, Sanjay Vishwakarma, Chunpu Wang, Y. C. Jhang, Maciek Chociej, Adam Crespi, Alex Thaman and Sujoy Ganguly, "PSP-HDRI+: A Synthetic Dataset Generator for Pre-Training of Human-Centric Computer Vision Models," 2022, ArXiv abs/2207.05025



(그림 2) 어뷰징 행동 학습 데이터 세트 구축 결과

각성도 및 긍·부정도를 적용한 멀티모달 감정인식 모델

신유담*, 김범수**, 서상현*
 중앙대학교 예술공학부*

중앙대학교 예술공학대학원 응용예술공학과**

e-mail : yudam616@cau.ac.kr, bumsookim@cau.ac.kr, sanghyun@cau.ac.kr

Adapting the Arousal-Valence Model for Multimodal Emotion Recognition with Incomplete Data

Yudam Shin*, Bumsoo Kim**, Sanghyun Seo*

The College of Art and Technology, Chung-Ang University*

Applied Art and Technology, Chung-Ang University**

요약

멀티모달 감정인식은 상호 보완적인 정보를 통합하여 인간의 감정 표현을 정밀하게 분석할 수 있다. 각성도와 긍·부정도의 두 축으로 표현하는 Arousal-Valence 모델은 기존 이론의 복잡한 감정에 대한 한계를 보완하는 반면 영상과 같은 특정 데이터 형식에서만 주로 활용되며, 다양한 데이터 형식에 적용하는 데 어려움이 있다. 본 논문에서는 오디오와 텍스트를 사용하여 각성도와 긍·부정도를 기반으로 감정인식을 수행하며, 불완전한 데이터 상황에서도 강건한 멀티모달 감정인식 모델을 제안한다. 오디오와 텍스트의 각성도, 긍·부정도를 예측하는 사전 학습된 모델을 결합하여 Arousal-Valence 모델을 다양한 데이터 형식에 적용할 수 있도록 하였다.

1. 서론

최근 인공지능의 발전과 함께 감정인식은 활발한 연구 주제가 되고 있다. 기본 감정 이론을 바탕으로 6가지 감정을 분류하거나 연속적인 감정의 척도를 인식하는 다양한 작업이 이루어지고 있다[1]. 음성, 뇌파, 이미지 등 사용할 수 있는 데이터의 형식이 다양해짐에 따라 멀티모달 감정인식에 대한 연구도 증가하고 있다.

멀티모달 감정인식 연구의 가장 큰 문제점 중 하나는 특정 모달리티가 누락되는 상황에서의 견고성이다. 대부분의 모델은 모든 모달리티가 존재한다는 가정 하에 설계되어, 모달리티가 누락된 상황에서 성능이 저하되는 한계가 존재한다[2]. 또한 한국어의 특성상 동일한 표현도 문맥이나 목소리의 크기에 따라 의미가 달라질 수 있어, 기존의 언어적 요인에 초점을 맞춘 모델로는 감정인식의 정확도에 한계가 존재한다. 특히 화자의 목소리 억양이나 속도와 같은 비언어적 요소는 감정 상태와 깊은 관련이 있으며, 감정인식 과정에서 고려되어야 하는 요소이다[3].

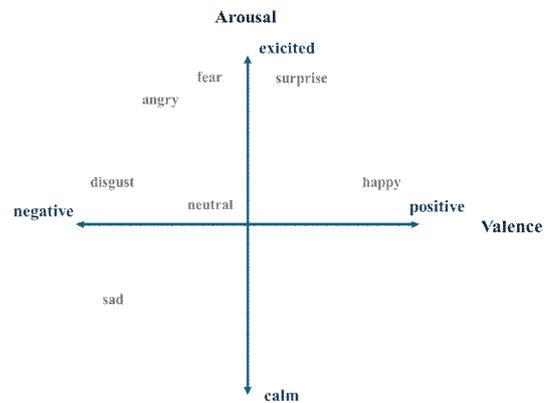
이러한 문제를 해결하기 위해 본 논문에서는 각성도와 긍·부정도를 적용하여 불완전한 데이터에서도 견고한 감정인식 모델을 제안한다.

2. 관련 연구

2.1 Arousal-Valence model

Arousal-Valence 모델은 각성도와 긍·부정도를 지표로 감정을 표현한다[4]. 각성도는 흥분의 정도를 나타내며, 낮은 값은 차분한 상태를, 높은 값은 흥분된 상태를 의미한

다. 긍·부정도는 감정이 내포하는 긍정적 또는 부정적인 정도를 나타낸다. 이 모델은 감정의 수치화를 가능하게 하여, 연속적인 평면을 통해 감정의 세밀한 표현을 시도한다.



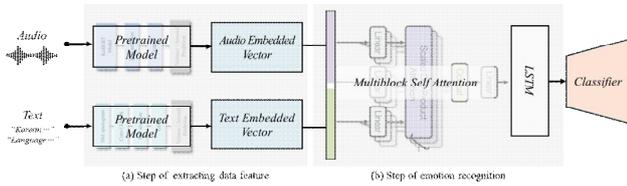
(그림 1) Arousal-Valence 모델 감정 예시

2.2 ETRI KEMDy20

본 연구에 사용된 데이터는 2020년 한국 다중 모달 감정 데이터셋(KEMDy20)이다[5]. 이 데이터셋은 음성, 텍스트, 생체 신호의 세 가지 모달리티를 포함하고 있다. 일반인을 대상으로 6개의 주제에 대한 자유 발화 과정을 통해 수집되었으며, 각 발화는 10명의 평가자에 의해 감정, 각성도, 긍·부정도로 평가되었다. 평가 결과를 바탕으로 7가지 감정 중 최다 선택된 감정과 각성도 및 긍·부정도의 평균값이 레이블로 사용되었다.

3. 각성도 및 긍·부정도 반영 멀티모달 감정인식

본 논문에서는 각성도 및 긍·부정도를 반영한 멀티모달 감정인식을 세 가지의 블록으로 구성하였다. 그림 2는 멀티모달 감정인식 모델의 전체 구조를 보여준다.



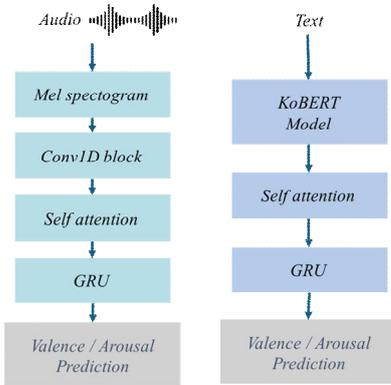
(그림 2) 각성도 및 긍·부정도 반영 멀티모달 감정인식 모델 구조

3.1 오디오 기반 예측 모델

오디오 데이터는 멜 스펙트로그램으로 변환되었으며, 특징 추출을 위해 1차원 컨볼루션 신경망을 사용하였다. 트랜스포머 모델의 인코더 구조를 사용한 self-attention 기법을 사용하여 오디오의 각성도와 긍·부정도를 예측하였다[6]. 예측 모델의 파라미터를 기반으로 오디오 임베딩 레이어와 오디오 인코더를 사용하여 오디오 데이터의 특징 추출을 진행하였다.

3.2 텍스트 기반 예측 모델

텍스트 데이터는 KoBERT[7]를 사용하여 임베딩되었으며, 텍스트의 각성도와 긍·부정도를 예측하기 위해 트랜스포머 모델의 인코더 구조를 활용한 self-attention 기법이 적용되었다. 텍스트의 특징을 추출하는 과정에서는 텍스트 기반 예측 모델의 예측 레이어가 생략되었다.



(그림 3) 각성도 및 긍·부정도 예측 모델 구조

3.3 감정인식

특징 추출 과정에서 오디오 임베딩 벡터와 텍스트 임베딩 벡터를 병합한 후, Multi-head attention 기법을 사용하여 다양한 맥락적 특징을 추출하였다. 이후 LSTM 네트워크를 활용하여 최종적으로 감정을 인식했다.

4. 결과

단일 모달 기반 각성도 및 긍·부정도 예측 모델을 먼저 훈련한 후, 이를 바탕으로 멀티모달 감정인식 모델을 훈련하였다. 모델 1은 오디오 기반 각성도와 텍스트 기반 긍·부정도 예측 모델을 사용하였고, 모델 2는 오디오 기반 긍·부정도와 텍스트 기반 각성도 예측 모델을 사용했다. 모델 3은 오디오 기반 각성도 및 긍·부정도 예측 모델, 모

델 4는 텍스트 기반 각성도 및 긍·부정도 예측 모델을 활용하였다. 훈련이 완료된 후 모델의 성능은 정확도와 F1 점수를 통해 정량적으로 평가되었다.

(표 1) 멀티모달 감정인식 모델 평가 결과

Model number	Used prediction model	accuracy	F1-Score
1	audio-based arousal prediction model	90.4%	88.8%
	text-based valence prediction model		
2	audio-based valence prediction model	87.8%	86.6%
	text-based arousal prediction model		
3	audio-based arousal prediction model	90.0%	88.6%
	audio-based valence prediction model		
4	text-based valence prediction model	81.9%	81.5%
	text-based arousal prediction model		

표 1은 멀티모달 감정인식 모델의 평가 결과를 보여준다. 모델 1은 90.4%의 정확도와 88.8%의 F1 점수를 통해 상대적으로 가장 우수한 성능을 보였다.

5. 결론

본 논문에서는 멀티모달 감정인식에 각성도와 긍·부정도를 적용하였다. 오디오와 텍스트의 각성도 및 긍·부정도 예측 모델의 다양한 조합의 실험함으로써, 각성도와 긍·부정도를 오디오와 텍스트에 적용하는 효과적인 방법을 확인할 수 있었다. 향후 연구에서는 영상과 생체 신호와 같은 다양한 형식의 데이터를 추가하여 더 세밀하고 정확한 감정인식이 가능하게 할 계획이다.

감사의 글

이 연구는 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0020632, 2022년 산업혁신인재성장지원사업)

참고문헌

- [1] P. Ekman, "Are there basic emotions?" 1992.
- [2] Lei Cai, Zhengyang Wang, Hongyang Gao, Ding-gan g Shen, and Shuiwang Ji, "Deep adversarial learning for multi-modality missing data comple-tion," in Proceedings of the 24th ACM SIGKDD in-ternational conference on knowledge discovery & data mining, 2018, pp. 1158 - 1166.
- [3] Anders, Silke et al. "Brain activity underlying emo-tional valence and arousal: a response-related fMRI stud y." Human brain mapping vol. 23,4 ,2004 : 200-9. doi:10.1002/hbm.20048
- [4] Russell, James A. "A circumplex model of affect," Journal of personality and social psychology 39.6, 1161, 1980.
- [5] K. J. Noh and H. Jeong, "KEMDy20," https://nanum.etri.re.kr/share/kjnoh/KEMDy20?lang=ko_KR
- [6] VASWANI, Ashish. "Attention is all you need." arXiv preprint arXiv:1706.03762, 2017.
- [7] SKTBrain "KoBERT" <https://github.com/SKTBrain/KoBERT>

이미지 증강을 이용한 자율 주행 2D 객체 감지 최적화: 심층 주행 데이터셋 분석 및 실시간 안드로이드 구현

우우르 에르셀릭
전남대학교, 인공지능융합학과
ugurercelik@hotmail.com

Optimizing 2D Object Detection for Autonomous Driving using Image Augmentation : Deep Drive Dataset Analysis and Real-Time Android Implementation

ERCELIK UGUR
CHONNAM NATIONAL UNIVERSITY , DEPARTMENT OF ARTIFICIAL
INTELLIGENCE CONVERGENCE

Abstract

Object Detection is fundamental component in the development of autonomous vehicles, enabling them to perceive and understand their surroundings in real time. In the pursuit of fully autonomous vehicles, real-time object detection remains a critical challenge for ensuring safe navigation in dynamic environments. Moreover, for autonomous vehicles, object detection is critical aspect in terms of enabling them to understand and navigate their environment. This paper mainly focuses on 2D Object detection where the goal is to identify and localize various objects such as cars, pedestrians, traffic signs, and obstacles in two-dimensional space. While doing this, we use various image augmentation methods to make the results suitable for the real world in object detection and to increase performance. We also test our results on the Mediapipe Android architecture.

1. Introduction

Autonomous vehicles (AVs) are rapidly advancing technologies poised to revolutionize transportation, promising to enhance road safety, improve traffic efficiency, and reduce human errors that contribute to accidents. One of the fundamental components enabling autonomous vehicles to navigate complex environments is object detection. Object detection, particularly in the 2D domain, involves identifying and localizing key objects such as vehicles, pedestrians, cyclists, traffic signs, and obstacles within the vehicle's surrounding environment using image data from cameras.

In this study, we use image augmentation methods to improve performance when performing 2D object recognition and detection on autonomous vehicles. Then, we test the model we obtain as a result of augmentation with Mediapipe Android Architecture on the real world in the Android environment. We use CODA Deep drive as a dataset.

2. Related Works

In the field of 2D Object Detection for Autonomous Driving, several research studies have been carried out. This section presents different research studies revolving around 2D Object Detection. Liv et al., [1] Effectively detecting road objects in

various environments would significantly improve drivingsafety for autonomous vehicles. However, small objects, low illumination, and blurred outline in images strongly limit the performance of current road object detection methods. To solve these problems, paper proposed a novel deep learning anchor-free approach based on CenterNet. The atrous spatial pyramid pooling (ASPP) was used to extract features from multiple scales to improve the detection performance while not increasing the computational cost and the number of parameters. Cao et al., [2] Effectively detecting road objects in various environments would significantly improve drivingsafety for autonomous vehicles. However, small objects, low illumination, and blurred outline in images strongly limit the performance of current road object detection methods. To solve these problems, proposed a novel deep learning anchor-free approach based on CenterNet. The atrous spatial pyramid pooling (ASPP) was used to extract features from multiple scales to improve the detection performance while not increasing the computational cost and the number of parameters.

3. Methodology

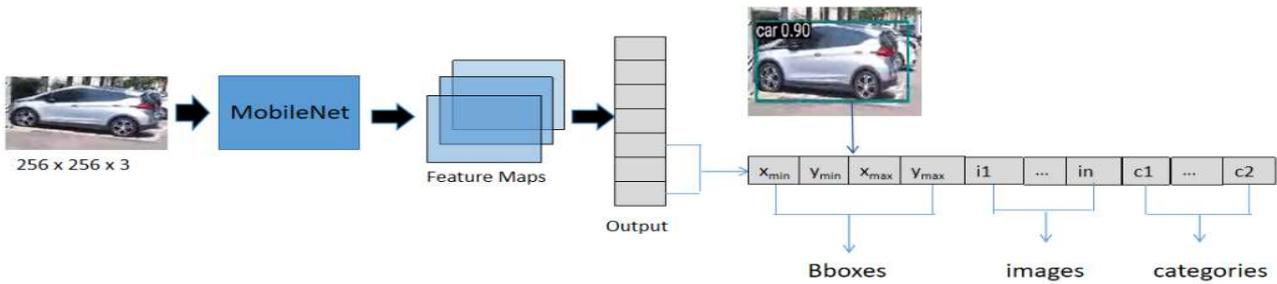


Figure 1 Architecture of Training Process

We will share the methods we used and how we achieved the results within the scope of this project. After training the augmentation or normal images we use in the train process through the mobilenet model, we obtain the feature maps of each image and determine the bounding box values, type values and classification results of the relevant image. In addition, we augment the image of the object, the bounding box values in the augmentation process and obtain the category values with the augmented results. The stages are as shown in the **Figure 1**

3.1 Without Augmentation

First, we train the CODA Deep Driveset dataset without the augmentation methods. The dataset consists of the train and validation images as well as the annotations of these images. The annotations json file consist of the image, categories and annotation values of the images. We trained the images and the annotations json file that we prepared in COCO Format for the images using the **MOBILENET_MULTI_AVG_I384 (Figure1)** architecture. In Training, Epoch is set to 30 and batch size is set to 32.

3.2 With Augmentation

In this part, we applied augmentation methods on the CODA Deep Drive Dataset. The augmentation methods we use are as follows:

- Horizontal Flip (p=0.5)
- Random Brightness Contrast (p=0.5)
- Rotate (limit = 15, p = 0.5)
- Random Rain (p=0.2)
- Random Snow (p = 0.2)

The train and test data sets we have before using the augmentation method and after augmentation are shown in **Table 1.**

	Train Set	Test Set
Normal Dataset	3884	1000
Augmented Dataset	12689	3253

Table 1 Train And Test Set Size After Augmentation

4. Results

We compared the results of the new images after the augmentation method with the images without using the

augmentation methods in **Table 2.** After these results, we tested the .tflite extension model we created on the Mediapipe Android Architecture to test our model on real world images and we were able to get the bounding box and object type of the relevant objects.

	Total Loss	Classification Loss	Box Loss	Model Loss	AP	AP50	AP75
Normal Dataset	0.95	0.65	0.0048	0.89	4.20	6.65	4.61
Augmented Dataset	0.53	0.29	0.0035	0.4647	0.121	0.196	0.128

Table 2 Evaluation Results

5. Conclusion and Future Works

In this paper , we proposed a 2D object recognition project for autonomous vehicles using the CODA Deep Drive Dataset and obtained the initial results. In this context, we evaluated how the Dataset can be detected under different conditions. Firstly, we used Normal Dataset and tested all Images. Secondly , we used augmentation methods for improving results. In future stages, we will try to develop a new model and comparing with other models.

Acknowledgement

This work was supported by the Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea government(MOTIE-P0024554) This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government(MSIT).

References

[1] G. Li, H. Xie, W. Yan, Y. Chang and X. Qu, "Detection of Road Objects With Small Appearance in Images for Autonomous Driving in Various Traffic Situations Using a Deep Learning Based Approach," in *IEEE Access*, vol. 8, pp. 211164-211172, 2020, doi: 10.1109/ACCESS.2020.3036620

[2] Y. Cao, C. Li, Y. Peng and H. Ru, "MCS-YOLO: A Multiscale Object Detection Method for Autonomous Driving Road Environment Recognition," in *IEEE Access*, vol. 11, pp. 22342-22354, 2023, doi: 10.1109/ACCESS.2023.3252021

Kano 모델 및 PCSI Index를 활용한 OTT 스포츠 중계 품질 개선 방안 연구

박미연¹⁾, 최은정²⁾, 이수철³⁾

- 1) 상명대학교 일반대학원 경영학과 석사과정, slsm8966@gmail.com
- 2) 상명대학교 경영학부 교수, e-mail : choiej@smu.ac.kr
- 3) 상명대학교 경영학부 조교수, e-mail : suchul.lee@smu.ac.kr, 교신저자

Analysis of OTT Sports Broadcasting Quality Improvement Using Kano Model and PCSI Index

Mi Yeon Park, Eun Jung Choi, Suchul Lee
Department of Business Administration, Sangmyung University

요 약

OTT (Over-The-Top) 서비스는 스포츠 중계의 새로운 유통 채널로 빠르게 자리 잡고 있으며, 기존의 전통적인 미디어 플랫폼을 대체하는 중요한 매체로 성장하고 있다. 이에 따라 소비자들의 시청 경험을 향상시키기 위한 다양한 품질 개선에 대한 연구가 필요한 시점이다. 따라서 본 연구는 OTT 스포츠 중계 서비스의 주요 품질 요인을 체계적으로 분석하고, 이를 통해 소비자 만족도를 높이기 위한 구체적인 개선 방안을 제시하는 것을 목적으로 한다. e-SERVQUAL과 WebQual 모델을 기반으로 스포츠 중계와 OTT 서비스의 선행연구를 각각 검토하여 OTT 스포츠 중계 서비스의 주요 품질 요인을 선정하였다. 이후, Kano 모델, 고객만족계수, PCSI Index 및 P-S Matrix를 활용하여 각 품질 요인의 중요성을 분석하고, 개선이 필요한 요인의 우선순위를 도출한다. 본 연구의 결과는 OTT 스포츠 중계 플랫폼 경쟁력 강화를 위한 참고 자료로 활용될 수 있을 것이며, 더 나아가 소비자의 만족도 증대에 기여할 수 있을 것으로 기대된다.

1. 서 론

OTT (Over-The-Top) 서비스는 인터넷을 통해 콘텐츠를 제공하는 미디어 플랫폼으로(윤재현·김한구, 2022), 스포츠 중계의 새로운 유통 채널로 빠르게 자리 잡고 있다. 스포츠 콘텐츠는 지속적인 경기 일정과 안정적인 수요로 인해 OTT 서비스의 킬러 콘텐츠로 주목받고 있으며(이승환·이상학, 2023; 조원호·장원용, 2024), 다양한 플랫폼이 중계권 확보 경쟁에 적극적으로 참여하고 있다(김원제, 2024). 이에 따라 OTT 서비스는 스포츠 중계 소비자의 인식을 고려한 서비스 품질 개선에 대한 연구가 필요한 시점이다. 이에 본 연구는 소비자들의 만족도 증대를 위해 OTT 스포츠 중계 만족도 요인을 파악하고, 소비자 인식에 따른 품질 요소의 개선 우선순위를 분석하여 우선적으로 관리가 필요한 요인을 판단하고 만족도를 높일 수 있는 전략을 제시하고자 한다.

2. 이론적 배경 및 관련 연구

스포츠 중계 품질 요인 도출을 위해 e-SERVQUAL과 WebQual 모델을 기반으로 다양한 연구를 검토한 결과, 영상 및 음향 품질, 중계진행자의 전문성, 정보 제공, 시청자 참여 등이 스포츠 전문 채널 중계방송 품질에 영향을 주는 주요 요인으로 나타났다(전찬수, 2021). OTT 서비스는 개인 맞춤형 추천, 멀티뷰 기능, 결제 방법의 편의성,

직관적인 인터페이스 등이 시청자의 만족도를 높이는 중요한 요인으로 나타났다(윤재현, 김한구, 2022; 현호석·이형석, 2022). 본 연구는 이문규(2002)의 e-SERVQUAL 모델을 기준으로 스포츠 중계 품질 요인과 OTT 서비스의 특성을 함께 고려하여, OTT 스포츠 중계 품질 요인을 정리하였다.

3. 연구 방법론 및 결과

본 연구는 OTT 스포츠 중계 품질 요인의 개선 우선순위를 분석하기 위해 국내 스포츠 OTT 플랫폼(티빙, 쿠팡플레이 등) 구독자를 대상으로 설문조사를 실시하였다. 총 120명의 응답을 분석에 활용하여, 제품 및 서비스 품질 요소를 고객 만족도에 따라 분류하는 Kano 모델을 적용하여(Kano et al., 1984) 표 1과 같이 품질 요인을 분류하였다.

표 1. 요인별 Kano 특성 분류 결과

요인	Kano 특성						특성 분류
	당연적 (M)	일원적 (O)	매력적 (A)	무관심 (I)	모순적 (R)	회악적 (Q)	
경기 영상의 품질	42	45	11	19	2	1	O
경기 오디오의 품질	34	54	14	15	2	1	O
중계진행자의 전문적인 해설	18	79	9	12	2	0	O

다양한 정보제공	17	36	36	27	0	4	O, A
문자 중계 서비스	13	5	26	64	10	2	I
쉬운 결제 방법	38	35	19	25	1	2	M
인터페이스 디자인	16	40	32	29	1	2	O
PIP 시청모드	11	34	41	32	1	1	A
시청자 상호작용 기능	6	12	24	74	2	2	I
개인 맞춤 서비스	7	20	42	45	5	1	I
시청자 의견 수용	29	34	26	28	0	3	O
다양한 참여 이벤트	7	17	50	43	1	2	A
지연 및 버퍼링 정도	29	72	1	9	3	6	O

매력적 특성은 충족되었을 때 만족, 충족되지 않아도 불만을 일으키지 않는 특성으로, ‘스포츠 경기와 관련된 다양한 정보 제공’, ‘PIP 시청모드’, ‘다양한 참여 이벤트’ 요인이 이에 해당된다. ‘쉬운 결제 방법’은 당연히 충족되어야 한다고 인식되는 당연적 특성으로 분류되었다. 또한, 충족 여부와 상관없이 무관심한 특성으로는 ‘문자 중계 서비스’, ‘시청자 상호작용 기능’, ‘개인 맞춤 서비스’가 해당되는 것으로 나타났다.

표 2. 만족도 분석 결과

요인	만족 계수 (S)	불만족 계수 (D)	현재 만족 수준 (L)	현재 만족 위치 (P)	PCSI index	PCSI 순위
경기 영상의 품질	0.479	-0.744	3.450	-0.388	0.867	4
경기 오디오의 품질	0.581	-0.752	3.517	-0.558	1.139	2
중계진행자의 전문적인 해설	0.746	-0.822	3.550	-0.708	1.454	1
다양한 정보제공	0.621	-0.457	3.467	-0.183	0.803	5
문자 중계 서비스	0.287	-0.167	3.125	0.287	0.000	13
쉬운 결제 방법	0.462	-0.624	3.583	-0.624	1.085	3
인터페이스 디자인	0.615	-0.479	3.442	-0.140	0.756	7
PIP 시청모드	0.636	-0.381	3.475	-0.141	0.777	6
시청자 상호작용 기능	0.310	-0.155	3.267	0.166	0.144	12
개인 맞춤 서비스	0.544	-0.237	3.433	0.019	0.525	8
시청자 의견 수용	0.513	-0.538	3.250	0.226	0.287	9
다양한 참여 이벤트	0.573	-0.205	3.225	0.403	0.170	11
지연 및 버퍼링 정도	0.658	-0.910	3.200	0.401	0.257	10

표 2와 같이 PCSI Index와 P-S Matrix를 활용하여 개선 우선순위를 도출하였다. 분석 결과, 중계진행자의 전문 해설과 경기 영상 및 오디오 품질이 시청자 만족도에 가장 큰 영향을 미치는 요인으로 파악되었다. 반면 문자 중계 서비스와 시청자 상호작용 기능은 중요도가 상대적으로 낮아, 안정성 유지에 중점을 두는 것이 적절한 것으로 확인되었다. PCSI Index 분석을 통해 개선 우선순위가 높

은 요인으로는 중계진행자의 전문성, 경기 영상 및 오디오 품질이 도출되었다.

4. 결론

OTT 스포츠 중계 서비스의 경쟁력 강화를 위해 시청자 만족도에 큰 영향을 미치는 중계진행자의 전문성, 경기 영상 및 오디오 품질을 우선적으로 개선하는 것이 필요하다. 이러한 개선을 통해 시청 경험을 향상시킴으로써 플랫폼의 경쟁력과 소비자 충성도를 높일 수 있을 것으로 기대된다. 본 연구는 OTT 스포츠 중계 품질 개선을 위한 방향성을 제시하였으며, 향후 플랫폼 운영 전략 및 마케팅 전략 수립에 유용한 참고 자료로 활용될 수 있을 것이다.

참고 문헌

[1] 김원제, "OTT 시대 스포츠 중계와 보편적 시청권," 방송문화, 제428호, 2024, pp. 61-71.

[2] 윤재현, 김한구, "OTT 서비스 플랫폼 특성이 이용 만족을 통해 구독 유지 의도에 미치는 영향: 지각된 진단성의 조절효과를 중심으로," 경영연구, 제37권, 2022, pp. 25-40.

[3] 이문규, "e-SERVQUAL-인터넷 서비스 품질의 소비자 평가 측정 도구," 마케팅연구, 제17권 제1호, 2002, pp. 73-95.

[4] 이승환, 이상학, "OTT 스포츠 방송 서비스 품질 특성이 고객만족도와 행동의도에 미치는 영향: 스포츠 콘텐츠 관여도의 조절효과를 중심으로," 한국스포츠산업경영학회지, 제28권 제6호, 2023, pp. 1-18.

[5] 전찬수, "델파이 분석을 이용한 스포츠전문채널 중계 방송품질 평가 요인 추출 및 개선 우선순위 분석: Kano 모델과 PCSI 지수 활용," 한국스포츠학회, 제19권 제2호, 2021, pp. 159-169.

[6] 조원호, 장원용, "OTT 스포츠 중계방송 콘텐츠 특성과 시청 행동 연구," 한국엔터테인먼트산업학회논문지, 제18권 제2호, 2024, pp. 31-40.

[7] 현호석, 이형석, "OTT 서비스 선택요인에 대한 계층 분석과정의 적용," 한국컴퓨터정보학회논문지, 제27권 제10호, 2022, pp. 245-254.

[8] Kano, N., Seraku, N., Takahashi, F., Tsuji, S., "Attractive Quality and Must Be Quality," The Journal of The Japanese Society for Quality Control, Vol. 14, No. 2, 1984, pp. 39-48.

고령자 대상 다차원 인지 훈련 시스템의 교수자 경험 분석

김태홍^{1,2}, 정영재^{1,3}, 장경미^{1,3}, 도준형^{1,3}

¹한국과학기술정보연구원 오픈XR플랫폼융합연구단, ²한국한의학연구원 한의약데이터부,

³한국한의학연구원 디지털임상연구부

e-mail : {thkim, jhdo, jangkm, dudwoj68}@kiom.re.kr

An Analysis of Instructor Experience in a Multidimensional Cognitive Training System for Older Adults

Taehong Kim^{1,2}, Youngjae Jeong^{1,3}, Kyoung-Mi Jang^{1,3}, Jun-Hyeong Do^{1,3}

¹Convergence Research Center for Open XR Platform, Korea Institute of Oriental

Medicine, ²KM Data Division, ³Digital Health Research Division

요약

초고령화 사회의 도래로 치매 예방과 고령자 케어가 중요한 사회적 과제로 부각되고 있다. 본 연구는 확장현실(XR) 기술을 활용한 온라인 인지 및 신체 훈련 시스템을 대상으로 교수자 경험을 평가하고, 시스템 개선 방안을 도출하였다. 민간 및 공공 분야에서 활동하는 작업치료사, 운동처방사, 간호사를 대상으로 시스템 사용 후 설문조사와 심층 인터뷰를 통해 데이터를 수집·분석하였다. 평가 기준은 교수-학습 상호작용, 사용성, XR 시스템의 특성을 중심으로 설정되었다. 분석된 연구의 결과는 교수자들이 원거리에서도 고령자와의 원활한 상호작용을 통해 인지 및 신체 훈련을 수행할 수 있는 다차원 인지 훈련 시스템의 개선에 적용될 예정이다. 나아가 고령 학습자들이 사회적 고립을 극복하고 지속적으로 인지 훈련을 받을 수 있는 환경을 조성하는 데 기여할 수 있을 것으로 기대한다.

1. 서론

전 세계적으로 고령화가 가속화됨에 따라 고령자 케어와 치매 예방이 중요한 사회적 과제로 부각되고 있다[1]. 이러한 문제를 해결하기 위해 확장현실(XR) 기술을 적용해 몰입감과 상호작용을 증대해 고령자의 인지 및 신체 기능을 효과적으로 향상할 수 있고, 원거리에서도 고령자들이 사회적 활동을 지속할 수 있는 고령자 대상 다차원 인지 훈련 시스템을 개발하였다[2].

기존의 VR을 포함한 IT 기반 인지 훈련 시스템은 단독 수행으로 인해 사회적 기능 훈련이 어렵고, 고령자의 사용성 한계로 인해 효과적인 교육이 제한되었다. 반면, XR 시스템은 이러한 문제를 해결할 수 있는 다차원 학습 방법과 협업 중심의 콘텐츠를 개발하여 교수자들이 학습자들에게 더욱 적합한 교육을 제공할 수 있다.

본 연구에서는 시스템의 핵심 사용자인 교수자와 학습자 중 교수자 그룹을 대상으로 사용자 경험을 평가·분석하고 시스템 개선을 위한 방안을 제시하고자 한다.

2. XR 기반 다차원 인지 훈련 시스템

XR 기반 다차원 인지 훈련 시스템은 고령자가 익숙한 환경에서 화면을 통해 교수자와 대화로 의사소통하며 수업에 참여하고, 필요에 따라 HMD를 통해 VR 혹은 AR 콘텐츠를 다른 학습자 및 교수자와 함께 수행할 수 있는 시스템이다[1].



(그림 1) XR 기반 다차원 인지훈련 시스템

2.1 교수자 시스템

교수자 시스템은 작업치료사와 운동처방사의 현업 업무 분석을 통해 설계되었으며, 학습자 분석, 기록 조회, 수업 콘텐츠 등의 수업 준비 절차와 학습자의 상태를 모니터링하고 안정적으로 수업을 운영할 수 있도록는 화상과 VR 모니터링, 기기 상태 및 생체신호 기반 안전 모니터링 등 수업 진행 기능으로 구성되어 있다. 이러한 기능을 통해 교수는 원거리에서 다수 학습자의 진행 상황을 실시간으로 파악하고, 개인화된 피드백을 제공할 수 있다.

3. 평가 방법

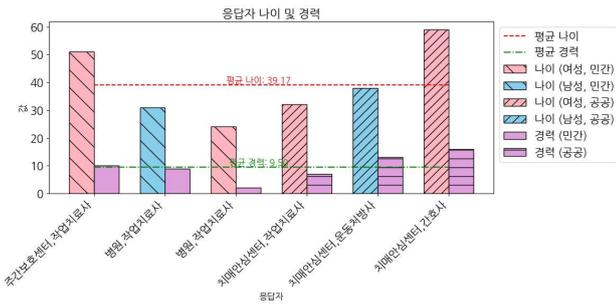
3.1 교수자 시스템 사용성 평가 기준

교수자 시스템의 사용성 평가 기준은 다음과 같이 구성하였다:

- 1) 교수-학습 평가 기준: Schueckle & Shuell의 교육용 소프트웨어 평가 준거와 한국학술정보연구원의 교수-학습 인증 기준[3].
- 2) 사용성 평가 기준: ISO 9241-11에서 제시하는 효용성, 효과성, 만족도.
- 3) 시스템 특성 평가 기준: XR 인지 훈련 콘텐츠에서 고려해야 할 인지, 신체, 사회 기능의 증대를 위한 디바이스 관련 기준[1]

4. 평가 결과 및 고찰

4.1 평가 대상

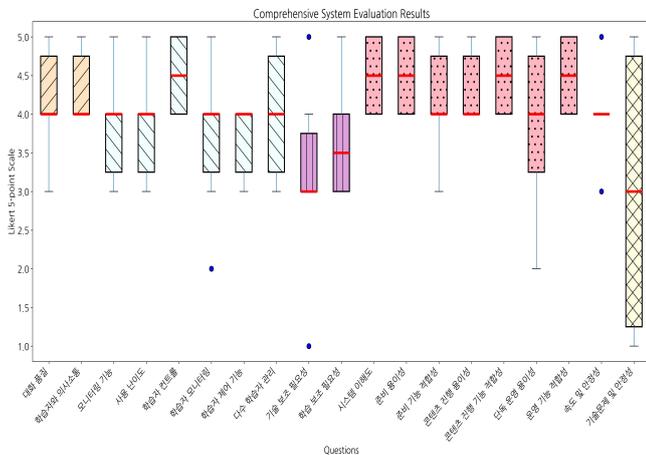


(그림 2) 교수자 사용성 평가 대상자 특성

교수자 시스템의 사용성 평가에는 민간 주간보호센터 및 병원, 공공 치매안심센터의 작업치료사, 운동처방사, 간호사 등 총 6명이 참여하였다. 평균 경력은 9.59년, 평균 나이는 39.17세이며, 모두 IT 기기(VR, 태블릿, 로봇)를 활용한 고령자 대상 교육 경험이 있었다. 객관적인 평가를 위해 대상자들은 시스템에 대한 사전 교육 후, 연구원 2인이 학습자 역할을 수행하는 환경에서 시스템을 단독으로 평가하고 설문 및 인터뷰를 진행하였다.

4.1 평가 결과

평가 대상자들은 기존 IT 기기 기반 교육에서 고령자의 기기 조작 미숙, 의사소통의 어려움, 성취도 파악 및 모니터링의 어려움, 개별 지도 불가 등을 문제점으로 인식하고 있었다. 그러나 XR 시스템에 대해서는 의사소통과 성취도



(그림 3) 교수자 사용성 평가 결과

파악 및 모니터링 기능이 우수하다고 평가하였다. 그림 3은 평가 결과를 그룹화한 것으로, 의사소통, 학습자 모니터링 기능, 보조자 필요성, 수업 준비 및 운영 기능, 안정성을 색상으로 구분하여 나타냈다.

- 1) 교수-학습 평가 기준: 교수자를 위한 고려사항 중 사회적 상호작용 및 교수 전략의 상호작용성, 사용자 운영법, 결과 기록 등은 의사소통 및 대화 품질, 수업 준비 및 운영 기능으로 평가되었으며, 전반적으로 가장 우수한 특성으로 나타났다.
- 2) 사용성 평가 기준: 사용 난이도, 학습자 모니터링, 기술 및 학습 보조 필요성에 대해 평가되었으며, 1)의 평가보다는 다소 낮게 나타났다. 특히 보조자의 필요성에 대해서는 평균 3~3.5점으로 수업 보조자가 필요한 것으로 나타났다.
- 3) 시스템 특성 평가 기준: 콘텐츠 진행의 용이성, 기능의 적합성, 속도 및 안정성에 대해 평가되었으며, 안정성을 제외하고 우수한 것으로 나타났다.

5. 결론 및 제언

XR 시스템은 직관적인 인터페이스와 실시간 의사소통 및 모니터링 기능에 대해 높은 만족도를 보였다. 그러나 시스템의 응답 속도와 안정성에서 개선이 필요하다는 점이 도출되었다. 또한 본 평가에서는 사회적 상호작용 및 교수전략의 상호작용성을 평가할 수 있는 기준인 가상환경에서의 공존감이 평가 기준에서 제외되었다.

평가 결과를 바탕으로 도출된 개선 사항은 고령자 학습자의 교육 효과를 극대화하고, 교수자들이 효과적으로 수업을 진행할 수 있는 시스템 개발에 기여할 수 있을 것으로 기대한다.

ACKNOWLEDGMENTS

이 논문은 2024년 정부(과학기술정보통신부)의 재원으로 국가과학기술연구회 오픈XR(OXR) 융합연구단사업(No. CRC21011)의 지원을 받아 수행된 연구임

참고 문헌

- [1] 김태홍, 김중일, 서정우, 도준형. (2022). Extended Reality 기반 고령자 대상 인지·운동 기능 훈련 워크벤치 개발 및 콘텐츠 설계 제안. 한국정보통신학회논문지, 26(11), 1755-1758.
- [2] 김태홍, 박주희, 김정훈, 도준형. (2022-07-10). 고령 사용자 경험 향상을 위한 신체 특성 기반 XR 보조 기술. 한국컴퓨터그래픽스학회 학술대회, 서울.
- [3] 이은원, 이수정. 웹 기반 교육용 소프트웨어의 사용성 평가기준 연구. 과학교육논총. 2004 Dec;17:161-70.

심층 강화 학습(DRL)을 이용한 동적 환경에서의 데이터 캐시 최적화

권민구, 송민석

인하대학교 컴퓨터공학과

e-mail : kwonalsrn@inha.edu, mssong@inha.ac.kr

Data Cache Optimization in Dynamic Environments Using Deep Reinforcement Learning (DRL)

Mingoo Kwon, Minseok Song

Department of Electrical and Computer Engineering, Inha University

요약

본 연구는 동적인 환경에서 파일 캐시 문제를 최적화하기 위해 Proximal Policy Optimization (PPO) 기반의 심층 강화학습(DRL) 모델을 제안한다. 기존의 탐욕적 알고리즘은 파일 접근률과 워크로드의 변화 등 동적 환경을 충분히 고려하지 못한다. 제안된 PPO 기반 알고리즘은 동적 환경과의 상호작용을 통해 최적의 캐시 결정을 내린다. 실험 결과, PPO 기반 알고리즘은 복잡하고 변화가 많은 환경에서 전통적인 탐욕적 알고리즘보다 높은 캐시 적중률과 우수한 성능을 보였다. 이는 제안된 프레임워크가 실제 시나리오에서 파일 캐시 최적화에 효과적임을 입증한다.

1. 서론

파일 캐시는 대역폭 절감과 응답 속도 개선을 위해 중앙 서버의 파일을 엣지 서버에 저장하는 방법이다. 그러나 엣지 서버의 용량이 제한되어 있어, 어떤 파일을 캐시할지 효율적으로 선택해야 하는 한계가 있다. 본 연구는 이러한 한계를 해결하고자 PPO 기반 심층 강화 학습(DRL) 알고리즘을 제안하였으며, 이 알고리즘은 동적 환경에서도 캐시 상태와 조건에 유연하게 대응하여 기존 알고리즘보다 높은 효율성과 성능을 제공한다.

2. 관련 연구

Cui와 그 외 연구진은 협력적 엣지 캐싱, 에너지 효율 개선, 실시간 캐싱 적응 등 DRL 기반 Q-학습 기법을 사용해서 교체와 중단 비용을 고려한 실시간 캐싱 방법을 제안했다 [1]. Dao와 연구진은 다중 사용자 적응형 비트레이트 스트리밍에서 적중률과 콘텐츠 품질 간의 트레이드오프를 해결하기 위해, 캐시 결정과 품질 수준을 동적으로 선택하는 정책을 제안했다. 이 문제는 중첩된 다차원 0/1 배낭 문제로 변환되어 협력적 진이 학습 기반 유전 알고리즘으로 해결되었다 [2].

3. 시스템 모델 및 문제 공식화

파일 캐시 시스템에서는 N^{file} 개의 파일이 저장된다. 여기서 $i \in \{1, \dots, N^{file}\}$ 는 시간에 따라 변할 수 있는 접근 확률을 가진다. 파일 i 의 접근률은 p_i 라고 하고, 다양한 분포를 사용하여 생성할 수 있다.

접근률은 주기 $k = \{k \in \{1, \dots, N^{cycle}\}\}$ 마다 변동될 수 있다. 변동된 접근률은 $p_i(k)$ 라고 정의하고, 접근률의 변동은 그룹 g 안에서 그룹끼리 변동한다. 변동된 접근률을 기록한 데이터셋을 D 라고 한다. 주기 k 의 캐시 서버에 접속하는 요청 수는 N_k^{demand} 라고 한다.

제안된 동적인 환경에서의 파일 캐시 문제는 캐시 저장 한도 내에서 캐시 적중률을 최대화하는 것을 목표로 한다. 이진 변수인 X_i 는 파일 i 의 캐시 유무를 나타내고 캐시할 때 $X_i = 1$ 로 표현한다. 데이터 파일 i 의 크기를 S_i 라고 하고, S^{limit} 을 캐시 용량이라고 한다. 캐시 용량 한도 내에서 캐시 적중률을 최대화하기 위해 파일 캐시 문제는 다음과 같이 공식화된다.

$$\begin{aligned} & \text{Maximize} \sum_{i=1}^{N^{file}} \sum_{k=1}^{N^{cycle}} X_i p_i(k) N_k^{demand}, \\ & \text{subject to} \sum_{i=1}^{N^{file}} S_i X_i \leq S^{limit}, \end{aligned} \quad (1)$$

주기 k 마다 변동하는 접근률 예측이 어려워 동적인 환경에서는 탐욕적 방법이 효과적이지 않다. 그러므로 동적인 환경에서 학습 기반 접근법이 필요하다.

4. 알고리즘

제안된 DRL 알고리즘 중 하나인 PPO 알고리즘은 마르코프 결정 과정(MDP) 모델을 기반으로 관찰 공간, 행동 공간, 보상 모델을 정의한다 [3].

관찰 공간은 $S_t^{obj} = \{p_i^{init}, S_i^{file}, U_t^{cur}\}$ 로 표현되는데, 각 의미는 다음과 같다. p_i^{init} 는 데이터셋 D 에서 초기 접

근률을 도출하여, 전체 파일의 접근률을 나타내고, S_i^{file} 는 전체 파일의 용량을 S^{limit} 에 비례하여 나타낸다. U_t^{cur} 는 시간 단계 1부터 t 까지 에이전트가 선택한 파일의 용량 합계를 S^{limit} 에 상대적으로 나타낸다.

행동 공간은 1부터 N^{file} 까지 정의되며, 시간 단계 t 에서 선택된 행동 a_t 에 따라 파일에 대한 보상을 얻는다. 한번 선택된 파일은 마스킹 기법을 통해 다시 선택되지 않도록 한다 [4].

보상 모델은 시간 단계 t 에서 에이전트가 선택한 파일의 보상(r_t)를 의미하고, 다음과 같다.

$$r_t = \frac{\sum_{k=1}^{N^{cycle}} p_{a_t}(k) N_k^{demand}}{N^{cycle}} \quad (2)$$

제안된 DRL 알고리즘은 먼저 초기화 단계에서 관찰 공간, 행동 공간, 그리고 데이터 시스템 환경을 설정한다. 그런 다음, 학습 과정에서 각 에피소드마다 타임 스텝과 리셋 단계를 반복하며, 타임 스텝에서는 행동에 따른 파일 보상을 계산하고, 리셋 단계에서는 다양한 환경 학습을 위해 접근률과 파일 요청 워크로드를 조정한다. 이렇게 갱신된 관찰 공간과 보상을 바탕으로 정책을 업데이트한다. 학습이 완료되면, 결정 단계에서 학습된 PPO 모델을 사용해 최적의 파일 캐시 결정을 내리고, Optuna를 통해 최적의 하이퍼파라미터를 설정한다 [5].

5. 실험 결과

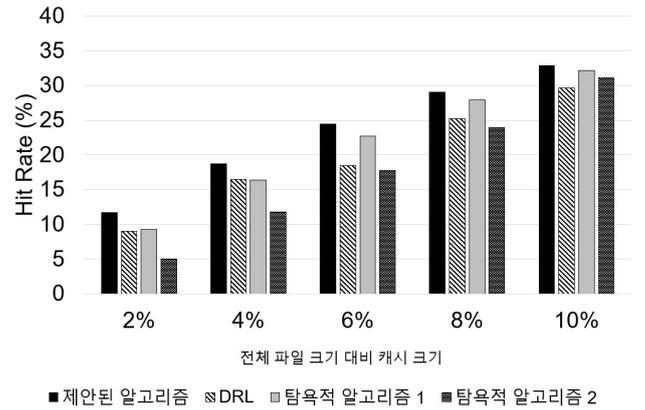
성능 평가를 위해 시뮬레이션을 진행했다. 파일 N^{file} 을 1000개로 설정하고, 파일의 용량을 2.25G에서 20.25G 안의 범위에서 무작위로 생성했다. 하루 동안의 접근률 변화를 조사하기 위해 N^{cycle} 은 24로 설정된다. 파일의 접근률은 Zipf 분포로 0.3에서 0.7까지 왜도 매개변수를 설정하고, 접근률 변화 그룹 g 는 10으로 설정했다. 각 실험에서 접근률 변화를 포착하기 위해 300개의 접근률 변화를 가진 데이터셋 D 를 생성했다. 한 개의 다른 DRL 알고리즘과 두 개의 탐욕적 알고리즘을 제안된 알고리즘과 비교했다. 비교 알고리즘은 다음과 같다.

DRL 알고리즘 : 제안된 알고리즘과 비교해 행동 공간 및 관찰 공간이 다른 상태에서 학습 후 파일 결정

탐욕적 알고리즘 1 : 접근률을 파일 크기에 비례해 높은 순으로 정렬해 순차적으로 결정

탐욕적 알고리즘 2 : 접근률이 큰 순으로 정렬해 순차적으로 결정

전체 파일 크기에 대한 캐시 크기에 따른 효과를 분석하기 위해 시뮬레이션을 수행했다. 전체 파일의 2%부터 10%로 설정하여, 제안된 알고리즘과 비교 알고리즘들을 비교해 그림 1로 나타냈다. 그 결과 캐시 적중률이 2.22%~57.04%(평균 18.22%)까지 더 높았고, 캐시 크기가 작을수록 더 좋은 결과를 나타냈다.



(그림 1) 다양한 크기별 캐시 적중률 비교

6. 결론

본 논문에서는 동적인 환경에서 파일 캐시 문제를 해결하기 위해 최적화된 강화학습 기반 알고리즘을 제안하고, 이를 기존 알고리즘과 비교 평가하였다. 실험 결과, 제안된 알고리즘은 비교 알고리즘보다 2.22%부터 57.04%, 평균 18.22% 더 높은 캐시 적중률을 달성했다.

7. 사사의 글

본 논문은 과학기술정보통신부에서 지원하는 정보통신기획평가원(IITP) 연구개발 사업(RS-2023-00261534)과 한국연구재단(NRF) 연구과제(NRF2022R1A2C1007237)의 지원을 받아 수행되었습니다.

참고문헌

- [1] L. Cui, E. Ni, Y. Zhou, Z. Wang, L. Zhang, J. Liu, Y. Xu, "Towards Real-Time Video Caching at Edge Servers: A Cost-Aware Deep Q-Learning Solution", IEEE Transactions on Multimedia, Vol. 25, pp. 302-314, November 2021.
- [2] N. Dao, D. Ngo, N. Dinh, T. Phan, N. Vo, S. Cho, and T. Braun, "Hit Ratio and Content Quality Tradeoff for Adaptive Bitrate Streaming in Edge Caching Systems", IEEE Systems Journal, vol. 15, pp. 5094-5097, December 2021.
- [3] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Kimov, "Proximal Policy Optimization Algorithms", arXiv preprint arXiv:1707.06347, July 2017.
- [4] C. Tang, C. Liu, W. Chen, S. You, "Implementing action mask in proximal policy optimization (PPO) algorithm", ICT Express, vol. 6, pp. 200-203, September 2020.
- [5] T. Akiba, S. Sano, T. Yanase, T. Ohta, M.Koyama, "Optuna: A Next-generation Hyperparameter Optimization Framework", In Proceedings of the CoRR, August 2019.

채용공고 데이터마이닝을 활용한 테크니컬 스킬 요구에 대한 분석

- 제조업 인공지능 관련 일자리 중심으로 -

윤고은, 이충권
 계명대학교 경영정보학과
 e-mail : haneul010731@gmail.com

A Study on the Technical Skill Requirements for AI-related Jobs in Manufacturing Industry

GoEun Youn, Choong Kwon Lee
 Department of MIS, Keimyung University

요약

인공지능 기술의 발전은 제조업을 포함한 다양한 산업에서 생산성 향상과 자동화의 가속화 등을 촉진하며, 이에 따라 관련 인재에 대한 수요도 급증하고 있다. 본 연구는 제조업에 속한 기업들이 제공하는 인공지능 관련 일자리에서 요구되는 테크니컬 스킬을 분석하기 위해 데이터마이닝 기법을 활용하여 채용공고 데이터를 수집하고 분석하였다. 본 연구의 결과는 제조업에 속한 기업들이 인공지능 관련 일자리에서 요구하는 기술역량을 구체적으로 파악하고, 이를 통해 구직자와 기업 모두에게 실질적인 가이드를 제공할 것으로 기대된다.

1. 서론

4차 산업혁명의 AI 기술 발전으로 기업들은 우수한 인재의 채용에 나서고 있지만, 채용시장에서의 경쟁은 더욱 치열해지고 있다. 본 연구는 AI 관련 일자리에 요구되는 테크니컬 스킬과 기업이 중점적으로 요구하는 역량을 분석하고자 한다.

2. 선행 연구

일자리에서 요구되는 지식과 스킬에 대한 연구는 오랫동안 지속적으로 연구되었다. 김소영(2022)은 비서직 채용공고 997건을 대상으로 그 내용을 분석하였다[1]. 임승주 등(2020)은 작업치료사 채용공고를 분석하였다[2]. 이진백과 이충권(2023)은 워크넷에 게재되는 취약계층을 대상으로 하는 137개의 채용공고를 수집하여 내용을 분석하였다[3]. 이처럼 채용공고를 활용한 일자리에 대한 연구의 경우 특정 직업에 대한 분석이 주로 이루어졌다. 그러나 특정 산업분야에서 인공지능 관련 직종에서 요구되는 테크니컬 스킬을 분석한 연구는 부족하다. 따라서 본 연구는 채용공고를 수집하여 제조업에서 제공하는 인공지능 관련 일자리에서 요구되는 테크니컬 스킬에 대해 분석하고자 한다.

3. 자료 수집

본 연구는 국내의 대표적인 잡포털 사이트인 JobKorea를 통해 제조업에 속한 기업들이 게재하는 인공지능 관련 채용공고 데이터를 2024년 6월 수집하였다. JobKorea 사이트에서 제공하는 분류 기준에 따라 제조·생산·화학업을

선택한 후에 검색어 “인공지능”을 입력하여 나타나는 채용공고 226개의 데이터를 수집하였다. 수집된 모든 데이터에서 특수기호, 공백, 불용어 등을 처리하여 N-gram, 연관분석 등의 기법을 활용하여 분석하였다.

4. 연구 결과

4.1 산업분야 별 요구 현황

표 1의 경력 항목의 경우, 하나의 채용공고에서 신입과 경력을 함께 모집하는 경우가 존재하여 신입과 경력을 하나씩으로 취급하였다. 산업분야 별 경력 요구 현황을 살펴본 결과에 따르면, 반도체/디스플레이/광학, 전기/전자제어 분야에서 경력 1~3년을 가장 많이 요구하고 있었다. 반면, 자동차/조선/철강/항공 분야에서는 경력무관이 가장 많이 나타나고 있었다.

표 1 산업분야 별 경력 요구 현황

분야/경력	경력 무관	경력 1년~3년	경력 4년~6년	경력 7년~9년	경력 10년이상	신입	총합
반도체	21	47	18	2	3	17	108
전기	15	28	7	0	0	10	60
자동차	14	7	11	4	5	1	42
화학	5	2	3	1	0	1	12
기계	4	0	1	0	4	0	22
섬유의류	2	0	5	0	0	1	17
생활화학	0	3	2	1	0	0	12
생활용품	2	0	1	0	0	2	12
농업	1	1	0	0	0	1	7

표 2에서 볼 수 있듯이 산업분야 별 지역 현황을 분석한 결과, 인공지능 관련 일자리의 채용공고는 산업 분야에 따라 특정 지역에 집중되는 경향이 있음을 확인할 수 있었다. 특히, 서울과 경기도 지역은 반도체/디스플레이/광학, 전기/전자제어 등 첨단 산업의 중심지로, 인공지능 기술 도입과 관련된 채용 수요가 매우 높았다. 이 결과는 제조업의 인공지능 활용이 특정 지역에 집중되어 진행되고 있음을 보여주며, 이는 지역 경제와 고용시장에도 중요한 영향을 미칠 수 있다.

표 2 산업분야 별 지역 현황

분야/급여	경기	서울	대전	인천	경북	경남	충남	대구	전북	부산	총합
반도체	44	33	4	4	5	0	2	2	1	0	95
전기	21	16	4	4	4	0	1	1	1	0	52
자동차	15	16	3	1	0	5	1	0	0	0	41
화학	4	4	2	0	0	0	0	0	0	1	11
기계	4	2	0	0	0	2	1	0	0	0	9
섬유의류	2	5	0	0	0	0	0	0	0	0	7
생활화학	0	6	0	0	0	0	0	0	0	0	6
생활용품	2	1	0	0	0	0	0	0	0	0	3
농업	1	1	0	0	0	0	0	0	0	0	2

4.2 직무내용 분석

연관분석을 통해 모집요강에서 나타내고 있는 테크니컬 스킬들의 연관성을 파악하였으며 support 값을 기준으로 상위 10개의 결과를 도출하였다. 표 3을 살펴보면 가장 높은 연관 규칙은 개발->AI로 support 값이 0.429로 개발이라는 테크니컬 스킬이 모집요강에 나타났을 경우, AI라는 테크니컬 스킬 또한 함께 나타나는 연관성을 확인할 수 있었다.

표 3 연관분석 상위 10개

Antecedent	Consequents	Support	Confidence	lift	fit
개발	AI	0.429	0.711	1.028	0.313
Python	AI	0.3	0.650	0.940	0.183
개발	Python	0.290	0.481	1.047	0.146
딥러닝	개발	0.268	0.740	1.229	0.244
머신러닝	AI	0.254	0.713	1.03	0.187
개발	머신러닝	0.254	0.422	1.182	0.127
딥러닝	AI	0.232	0.642	0.928	0.138
개발	C++	0.219	0.363	1.178	0.094
딥러닝	머신러닝	0.219	0.605	1.694	0.224
개발	Computer vision	0.214	0.356	1.284	0.098

Python->AI가 support 값이 0.3으로 Python이라는 테크니컬 스킬에 해당하는 프로그래밍 언어가 모집요강에서 나타났을 경우 AI라는 테크니컬 스킬 또한 함께 나타나는 확인할 수 있었다. Python은 인공지능 및 데이터 과학 분야에서 널리 사용되는 프로그래밍 언어로, 그 간결성과 강력한 라이브러리 생태계로 인해 많은 AI 프로젝트(머신러닝, 딥러닝, 데이터 분석 등)에서 주요 도구로 활용되기 때문에 인공지능 직무에서 자주 언급됨을 확인하였다.

5. 결론

본 연구는 제조업 분야의 인공지능(AI) 관련 일자리에서 요구되는 테크니컬 스킬을 파악하기 위해 데이터마이닝 기법을 활용하여 채용공고 데이터를 분석하였다. 이를 통해 몇 가지 중요한 결론을 도출할 수 있었다. 첫째, 반도체/디스플레이/광학 분야에서 AI 관련 기술에 대한 수요가 가장 높았으며, 이는 AI 기술이 반도체/디스플레이/광학 분야의 생산성 증대와 혁신에 중요한 역할을 하고 있음을 시사한다. 둘째, 경력 요구 현황을 분석한 결과, 인공지능 기술의 빠른 발전과 활용으로 인해 1~3년의 경력을 요구하는 경우가 많았으며, 이는 실무 경험을 통한 최신 기술 습득의 중요성을 반영한 것으로 보인다. 그러나 동시에 신입인력을 채용하여 새로운 기술에 대한 적응과 활용을 시도하려는 전략도 확인할 수 있었다. 셋째, 지역분석의 결과, 인공지능 관련 일자리의 채용은 서울과 경기도에 집중되는 경향이 뚜렷했다. 이는 수도권 지역의 기업들이 첨단산업의 중심지로서 AI 기술 도입에 있어서 인력 수요를 가장 많이 차지하고 있음을 보여준다. 이러한 지역적 집중은 해당 지역의 경제 및 고용 시장에 중요한 영향을 미칠 것으로 예상된다. 마지막으로, 채용공고에서 나타난 테크니컬 스킬의 연관성 분석을 통해 ‘개발’, ‘AI’, ‘Python’ 등의 스킬이 밀접하게 연관되어 있음을 확인하였다. 이는 인공지능 시스템의 구현에서 소프트웨어 개발의 중요성 및 Python과 같은 특정 프로그래밍 언어의 필요성을 잘 보여준다. 특히 AI 관련 직무에서는 개발과 AI, Python 등이 주요한 기술적 요구사항으로 자리 잡고 있었다. 이러한 결과는 제조업 분야에서 AI 기술이 어떻게 활용되고 있으며, 이에 따라 인재들이 준비해야 할 기술적 역량이 무엇인지를 명확히 제시해 준다. 본 연구는 AI 구직자들이 변화하는 채용패턴에 효과적으로 대응할 수 있도록 구체적인 가이드를 제공할 뿐만 아니라, 기업들이 필요한 기술적 역량을 명확히 정의함으로써 인재 채용의 전략 수립에 기여할 수 있을 것으로 기대된다.

참고 문헌

[1] 김소영. (2022). 비서직채용현황 기초분석: 온라인 채용공고를 중심으로. 비서·사무경영연구, 31(4), 5-37, 10.35605/jss.2022.12.31.4.5

[2] 임승주, 양민아, 한대성, 박지혁. (2020). 채용공고 분석을 통한 작업치료사 구직 시장에 대한 연구. 재활치료과학, 9(2), 39-53.

[3] 이진백, 이충권. (2023). 채용공고 분석을 통한 취약계층 일자리에 관한 연구. 로고스경영연구, 21(2), 231-248.

호텔 예약 플랫폼에서 평점 수렴 상황에서 소비자 의사 결정에 관한 실험 연구

오호희*, 고준**

전남대학교 디지털미래융합서비스협동과정 대학원생*, 전남대학교 경영학과 교수**
whx1122334@gmail.com*, kjoon@jnu.ac.kr**

An Experimental Study on Consumer Decision-Making in Converging Rating Situations on Hotel Booking Platforms

Haoxi Wu*, Joon Koh**

Interdisciplinary Program of Digital Future Convergence Service, Chonnam National University*
Department of Business Administration, Chonnam National University**

요 약

본 연구는 온라인 여행 플랫폼에서 평점 유사성이 소비자 의사 결정에 미치는 영향을 정보 압축 이론(TIC)의 관점에서 분석한다. 온라인 여행 시장의 성장과 함께 평점 시스템은 소비자 선택에서 중요한 역할을 하고 있으나, 평점이 유사한 상황에서는 정보의 차별성이 약화되어 의사 결정이 어려워진다. 본 연구는 가격 민감도, 구전효과(e-WOM), 쾌락주의 및 실용주의 성향, 그리고 AI 추천 시스템이 이러한 상황에서 소비자의 의사 결정 과정에 미치는 영향을 다루며, 특히 이들 변수 간의 상호작용을 규명하기 위해 실험적 설계를 적용하였다. 이를 통해 소비자의 의사 결정 과정에서 발생하는 복잡성을 완화할 수 있는 전략적 방안을 제시하고, 온라인 여행 플랫폼의 평점 및 추천 시스템 개선을 위한 실증적 근거를 제공하고자 한다.

1. 서론

최근 소비자들이 항공권과 호텔 예약에 있어서 온라인 플랫폼과 앱을 선호하는 경향이 뚜렷하게 나타나고 있으며, 이러한 변화는 관광 산업의 구조에 중대한 영향을 미치고 있다. Statista의 보고서에 따르면, 2023년 전 세계 온라인 여행 시장 규모는 약 7,560억 달러에 달하며, 2027년에는 1조 달러를 초과할 것으로 전망된다[1]. Expedia와 Booking.com과 같은 주요 플랫폼의 재무 보고서에 따르면, 2023년에는 활성 사용자 수와 예약 건수가 전년 대비 약 20% 증가했으며, 2024년에는 이 성장률이 두 배로 확대될 것으로 예상된다[2].

이와 같은 온라인 여행 플랫폼의 성장과 더불어, 평점 시스템은 소비자의 의사 결정에 중요한 영향을 미치는 요소로 자리 잡고 있다. 연구에 따르면, 약 80% 이상의 소비자가 호텔 예약 시 다른 사용자의 평점과 리뷰를 참고하며, 평점이 높은 호텔은 더 많은 노출과 예약 기회를 얻는다[3]. 예를 들어, Booking.com 데이터에 따르면, 평점 8.0 이상 호텔의 예약률은 6.0 이하 호텔보다 약 30% 더 높다[2].

그러나 평점이 유사할 경우, 시스템의 유효성이 저하되며 의사 결정 과정이 복잡해질 수 있다. 평점이 일정 범위에 집중되면 차별성이 약화되어 소비자는 구체적인 리뷰에 더 의존하게 된다[5, 6]. Watson 등[7]은 이를 정보 압축 이론(TIC)으로 설명하며, 평점이 유사할수록 정보 가치가 감소하고 의사 결정 비용이 증가한다고 지적했다.

기존 연구들은 평점 다양성을 높이고 다차원 평점 모델을

도입하여 정보 압축 문제를 해결하려 했으나, 일부 연구에서는 평점 유사성이 단순한 기술적 문제를 넘어서 플랫폼 사용자들의 소셜 네트워크와 행동에 의해 영향을 받는다는 점을 지적했다[8, 9]. 이에 따라 정보 압축 문제는 다차원적 요인에서 기인하며, 기술적 접근만으로는 해결하기 어렵다. 본 연구는 기존 연구와 차별화하여 다양한 상황을 설정한 실험을 통해 평점이 유사한 상황에서 소비자가 어떻게 행동하는지를 분석한다. 특히 가격 민감도, 구전효과(e-WOM), 쾌락주의 및 실용주의 성향이 의사 결정에 미치는 영향을 다루며[10, 11], Amazon에서 사용되는 AI 추천 시스템이 의사 결정 복잡성을 완화하는 데 어떤 역할을 하는지 탐구하였다[12].

따라서 본 연구는 평점이 동일하거나 유사한 상황에서 소비자의 선택에 영향을 미치는 요인들을 심층적으로 분석하고, 이를 통해 소비자의 의사 결정 과정에서 발생하는 복잡성과 불확실성을 어떻게 줄일 수 있는지를 탐구하고자 한다. 이러한 연구는 향후 온라인 여행 플랫폼에서 평점 및 추천 시스템의 개선을 위한 이론적 및 실증적 근거를 제공할 것이다.

2. 관련 연구

2.1 정보 압축 이론 (TIC)

정보 압축 이론(TIC)은 의사 결정 네트워크에서 판단 비용이 증가함에 따라 정보 시스템이 생성하는 정보가 일관성을 띠며 다양성과 차별성을 상실하는 현상을 설명한다[13]. 이는 정보가 의사 결정을 지원하는 과정에서 변화

가 미미해 유용성이 감소할 때 발생하며, 결과적으로 시장 효율성과 의사 결정의 질을 저하시키는 요인이 된다[7]. 온라인 여행 플랫폼에서 이러한 정보 압축 현상은 호텔 및 관광지의 평점에서 두드러지게 나타난다. 대부분의 호텔 평점이 만점에 가까워지면 소비자는 평점을 기준으로 명확한 구분을 하기가 어려워지고, 정보의 유용성이 감소한다[7]. 이는 평점 시스템의 의사 결정 지원 기능을 약화시키며, 선택의 유효성 및 시장 경쟁력에도 부정적인 영향을 미친다[5, 6].

3. 실험 설계

본 연구의 실험 설계는 정보 유사성과 다양한 추천 방식이 사용자 의사결정 비용에 미치는 영향을 탐구하는 데 주요 목적을 두고 있으며, 두 가지 실험으로 나누어 진행된다.

첫 번째 실험에서는, 평점과 가격이 유사하거나 다를 때 소비자의 의사결정 비용이 어떻게 변화하는지를 탐구한다. 평점과 가격이 모두 유사한 경우, 소비자는 정보 압축 현상으로 인해 결정을 내리기 어려워지며, 이로 인해 의사결정 비용이 유의미하게 증가할 것으로 예상된다. 평점이나 가격 중 하나만 차이가 날 경우에는 소비자가 좀 더 쉽게 결정을 내릴 수 있을 것으로 보인다.

두 번째 실험에서는 리뷰 수와 추천 시스템(AI 추천, XAI 추천)이 소비자 의사결정에 미치는 영향을 분석한다. 리뷰 수, AI 추천, XAI 추천 모두가 소비자의 의사결정 비용을 줄이는 효과가 있을 것으로 예상되며, 특히 XAI 추천은 소비자에게 더 많은 설명과 투명성을 제공함으로써 더 큰 신뢰를 얻고, 결과적으로 소비자의 선택 자신감을 높일 가능성이 크다. 또한, XAI 추천을 사용한 소비자의 선택 비율이 AI 추천이나 리뷰 수에만 의존한 경우보다 더 높을 것으로 예측된다.

이를 통해, 본 연구는 호텔 예약 플랫폼에서 평점과 가격의 유사성, 리뷰 및 추천 시스템의 도입이 소비자의 의사결정 과정에 미치는 영향을 규명하고자 하며, 이러한 결과는 플랫폼의 최적화와 소비자의 만족도를 높이는 방안을 제시하는 데 기여할 수 있을 것으로 기대된다.

참고 문헌

[1] Statista, "Online travel market size worldwide 2017-2028", Statista Research Department, 2024.
 [2] Statista, "Most well-known travel portals in the United States 2023", Statista Research Department, 2023.
 [3] Tripadvisor, "Online Reviews Remain a Trusted Source of Information When Booking Trips, Reveals New Research", Media Center, 2019.
 [4] I.E. Vermeulen and D. Seegers, "Tried and tested: The impact of online hotel reviews on consumer consideration", *Tourism Management*, No.30, 2009, pp.

123-127.

[5] J. Lee and L.G. Pee, "Exploring a Mental Scaling Effect on Online Consumer Ratings: Identifications of Shopping Behaviors and Information Processing Rules", *International Journal of Electronic Commerce Studies*, Vol.9, No.1, 2018, pp. 1-26.
 [6] L. Ren, B. Zhu and Z. Xu, "Continuous Exp Strategy for Consumer Preference Analysis Based on Online Ratings", *IEEE transactions on fuzzy systems*, Vol.30, 2022, pp. 2621-2633.
 [7] R.T. Watson, K. Plangger, L. Pitt and A. Tiwana, "A Theory of Information Compression: When Judgments Are Costly", *Information Systems Research*, Vol.34, No.3, 2023, pp. 1089-1108.
 [8] X. Chen, Y. Pan and B. Luo, "Research on power-law distribution of long-tail data and its application to tourism recommendation", *Industrial Management & Data Systems*, Vol.121, 2020, pp. 1268-1286.
 [9] S. Park and J.L. Nicolau, "Asymmetric effects of online consumer reviews", *Annals of Tourism Research*, Vol.50, 2015, pp. 67-83.
 [10] M.y. Lee, Y.K. Kim and A.E. Fairhurst, "Shopping value in online auctions: Their antecedents and outcomes", *Journal of Retailing & Consumer Services*, Vol.16, 2009, pp. 75-82.
 [11] J. Ren and J.V. Nickerson, "Arousal, valence, and volume: how the influence of online review characteristics differs with respect to utilitarian and hedonic products", *European Journal of Information Systems*, Vol.28, 2018, pp. 272-290.
 [12] G. Linden, B. Smith and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering", *IEEE Internet Computing*, Vol.7, 2003, pp. 76-80.
 [13] P. Berthon, L.F. Pitt and R.T. Watson, "From Genesis to Revelations: the technology diaspora", *Communications of the ACM*, Vol.51, No.12, 2008, pp. 151 - 154.

디지털전환 관련 IT인력에 관한 연구 - 동남아시아 4개국 채용공고 분석 -

이진백, 이충권
계명대학교 경영정보학과
e-mail : kevin.yi@kmu.ac.kr. cklee@kmu.ac.kr

A Study on the IT Workforce for Digital Transformation - Analyzing Job Ads in Four Countries in Southeast Asia -

Jin Baek Yi, Choong Kwon Lee
Department of MIS, Keimyung University

요약

본 연구는 동남아시아 4개국(인도네시아, 말레이시아, 필리핀, 싱가포르)을 대상으로 2023년 1월부터 2024년 2월까지 IT인력을 구하는 채용공고를 수집하였고, 그중 Digital Transformation(DT)과 관련이 있는 86,808개의 채용공고를 분석하였다. 연구의 결과는 IT-DT 구직자에게 요구되는 지식과 기술이 국가별로 서로 다르다는 결과를 도출하였다.

1. 서론

디지털 전환(Digital Transformation, DT)은 전 세계적으로 기업들이 경쟁력을 유지하고 성장하기 위해 필수적인 전략으로 자리 잡고 있다. 특히, 동남아시아 지역은 빠르게 성장하는 경제와 기술 발전으로 인해 디지털 전환의 중요성이 더욱 부각되고 있다. 그러나 이 지역 내에서 일어나는 디지털 전환의 속도와 방향은 국가별로 상이하며, 이는 각국의 경제상황, 기술인프라, 교육체계, 산업구조 등 여러 요인에 의해 영향을 받는다.

본 연구는 동남아시아의 주요 4개국(인도네시아, 말레이시아, 필리핀, 싱가포르)을 대상으로 IT 인력을 구하는 채용공고를 수집하여, 이 중 디지털 전환과 관련된 채용공고를 별도로 추출하여 집중 분석하였다. 연구의 목적은 글로벌화된 디지털 경제 속에서 각국이 어떻게 서로 다른 접근 방식으로 디지털 전환을 추구하고 있으며, 각 국가별로 IT 인력의 요구사항이 어떻게 달라지는지를 파악하는 데 있다. 이를 통해 동남아시아의 디지털 전환을 촉진하고, 해당 지역에서 IT-DT 인재의 수요와 공급을 최적화하는 데 기여하고자 한다.

2. 이론적 배경

채용공고를 수집하여 분석하는 연구는 오랫동안 지속되었다. 연구의 방향은 경제적 측면에서 국가나 도시, 산업이나 기술의 관점에서 인력의 수요와 동향을 파악하고자 하는 시도가 있었다. 또한, 특정 직업에 대한 채용공고의 내용을 분석하여 해당 직업에서 필요로 하는 지식과 스킬을 파악하려는 연구가 다수 있었다(예, Acemoglu, et al., 2022). 특히, IT와 관련된 직업에 대한 연구가 많았는데, 웹개발자, 데이터분석가, 그리고 IT관리자 등과 같은 연구가 있었다. 또한 4차 산업혁명과 관련하여 빅데이터나 인공지능과 관련된 일자리에 요구되는 지식과 스킬을 파악하기 위하여 채용공고를 수집하여 분석을 시도한 연구도 있었다(예, Alekseeva, et al., 2021; Gilli 등, 2023). 본 연구는 채용공고의 내용을 분석하여 동남아

시아 국가들에서 IT-DT 인력에게 요구되는 지식과 기술을 이해하고, 국가별로 다르게 요구되는 지식과 스킬이 무엇인지 살펴보고자 한다.

3. 데이터 수집

본 연구는 2023년 1월부터 2024년 2월까지 동남아시아의 대표적인 잡포털 사이트(<https://www.jobstreet.com/>)로 부터 크롤링을 통하여 210,276건의 채용공고를 수집하였다. 수집된 채용공고 데이터는 잡포털사이트를 통해 IT인력 채용공고에 나타난 텍스트의 전문이 그대로 수집되었으므로, 분석과 관련이 없는 특수기호나 제어문자 등이 다수 포함되어 있었다. 특수문자는 제거하였고, 제어문자의 경우 일반문자와 결합하여 실제 분석하고자 하는 용어의 의미를 왜곡시킬 가능성이 있기 때문에 해당 문자들에 대해서는 함수를 사용하여 데이터를 정제하였다. 기존의 연구(송영근 등, 2022; Gilli 등, 2023)를 조사하고 분석하여 DT와 관련된 단어들을 나열하고 이 단어들을 포함하고 있는 채용공고들만을 추출하였다. 표 1과 같이 수집된 210,276개의 IT 채용공고들로부터 DT 관련 단어들을 포함하고 있는 IT-DT 채용공고 86,808개가 추출되었다. IT-DT 채용공고는 41.3%로 상당한 숫자의 IT 채용공고들이 DT 단어들을 포함하고 있는 것으로 확인되었다.

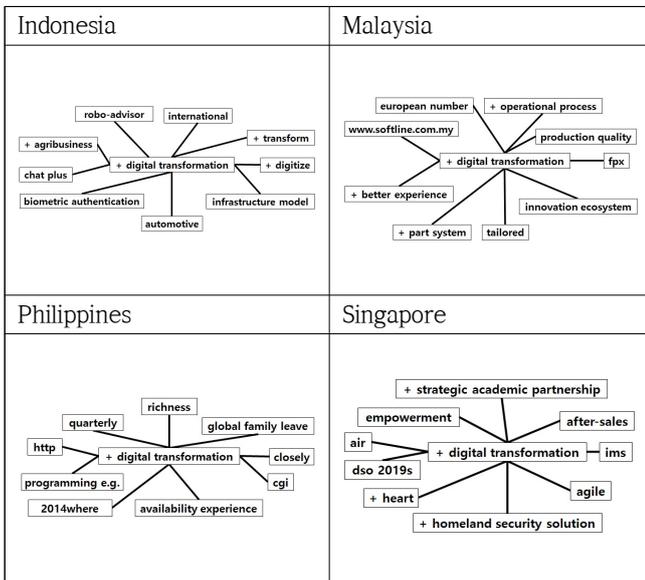
(표 1) Digital Transformation 채용공고

Country	Number of Job ads (IT)	Number of Job ads (IT-DT)	IT에 대한 IT-DT 비율(%)
Indonesia	23,029	5,635	24.5
Malaysia	36,674	13,475	36.7
Philippines	66,347	27,016	40.7
Singapore	84,226	40,682	48.3
합계	210,276	86,808	41.3

4. 분석결과

국가별로 IT인력 채용공고에서 빈도의 상위를 차지하고 있는 용어에 대해 해당 용어를 중심으로 관련된 단어들을 파악하는 Text Link Analysis(TLA)를 수행하였다. 이 기법은 분석의 대상이 되는 용어를 중심으로 연결된 용어들 간의 링크가 국가별로 상이하게 나타나는지를 확인할 수 있게 해준다. 표 2는 빈도 수 상위에 속한 용어들 중에서 DT에 대한 TLA를 수행한 결과이다. 국가별로 TLA에서 나타난 바와 같이 동일한 DT라는 용어에 대해서 링크되어 나타나는 용어들이 서로 다른 것으로 나타났다. 이것은 4개 국가들 간에 IT인력에 대한 수요의 조건이 반영된 것으로 분석된다.

(표 2) Information TLA



국가별 TLA 분석에서 나타나듯이 국가별로 DT라는 용어를 중심으로 각기 다른 용어가 링크 되어 있는 것을 확인 할 수 있다. 인도네시아의 경우 agribusines 라는 용어와 robo-advisor 라는 용어가 링크 되는 특징을 보였으며, 말레이시아의 경우 tailored, part system 등의 용어가 링크되는 것으로 DT와 관련된 특화된 산업 보다는 관련된 산업을 지원하는 IT인력에 대한 채용공고가 게재되는 것으로 파악할 수 있다. 필리핀의 경우에도 DT와 관련된 개발, 프로젝트를 수행하는 인력에 대한 용어가 링크되는 것으로 나타났다. 싱가포르의 경우 DT라는 용어가 다른 동남아 3개국과 달리 homeland security solution, strategic academic partnership 등과 같이 실행 범위가 큰 분야를 나타내는 용어와 agile, after-sales 등 개발 방법과 프로세스를 나타내는 용어가 같이 나타나는 것을 확인할 수 있다.

동남아 4개 국가의 IT-DT 채용공고에서 artificial intelligence 용어에 대해서 TLA를 수행해 본 결과 표 8과 같이 각 국가별로 artificial intelligence 용어에 대해서 관련된 용어들이 차이점을 보이는 것을 확인할 수 있었다. 이를 통해 각 국가별로 채용공고에서 구직자에게 요구하는 구체적인 내용은 상이하다는 것을 알 수 있다.

(표 8) artificial intelligence 용어에 대한 국가별 TLA 분석

Country	artificial intelligence TLA 결과
Indonesia	talented leader, Sunday, chatbot, analytical model, biometric authentication, chat plus, api management, robo-advisor, laptop
Malaysia	high-performance, brick-and-mortar business, next-generation, intelligent, device operating system solution, 3-3a, 2019s leader, recruitment service, chosen company
Philippines	departmential initiative, workforce community, training data, thales, mission, prestigious company, on matural language, value diversity, americas
Singapore	repeatable platform, security solution, exciting stage, homeland security solution, aws microservices, research outcome, air, stratigic academic partnership, professional programming

5. 시사점과 결론

동남아 4개 국가는 지리적으로는 가깝지만, 각 국가의 경제발전 단계와 산업이 서로 다르고, 종교, 문화, 생활방식 등에 있어서도 차이가 존재한다. 차별화된 부분의 용어들로 이는 각 국가의 경제 발달 및 산업의 상황과 밀접한 관계가 있다고 볼 수 있다. TLA 분석을 통해 동일한 용어를 중심으로 연관된 용어들이 각 국가별로 상이한 부분을 분석할 수 있다. 이러한 텍스트마이닝 기법분석을 통해 인도네시아와 필리핀의 경우 글로벌 company의 system 개발 및 유지보수 업무와 관련이 있는 용어 및 토픽을 나타내고 있으며, 말레이시아는 인도네시아 채용공고에서 요구되는 것과 유사하지만, 보다 구체적인 기술이 채용공고에 나타나는 것을 확인할 수 있었다. 싱가포르에서는 tiktok, video 등 다른 동남아 3개국과는 차별화된 IT기술 및 업종에 대한 인력의 채용을 원한다는 것을 확인할 수 있었다. 향후 연구는 각 국가별 차이가 채용공고에 어떻게 반영되는지 확인하고, 각 국가별 채용공고의 등장하는 용어와 분석에 있어서 다양한 기법들을 적용하여 의미있는 결과를 도출할 것으로 기대된다.

참고문헌

[1] 송영근, 박안선, 심진보. (2022). 디지털 전환의 개념과 디지털 전환의 범위. 한국전자통신연구원.
 [2] Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. Journal of Labor Economics, 40(S1), S293-S340.
 [3] Alekseeva, L., Azar, J., Gine, M., Samila, S., & Taska, B. (2021). The demand for AI skills in the labor market. Labour economics, 71, 102002.
 [4] Gilli, K., Nippa, M., & Knappstein, M. (2023). Leadership competencies for digital transformation: An exploratory content analysis of job advertisements. German Journal of Human Resource Management, 37(1), 50-75.

유럽 디자인 보호규정 개정이 디지털 디자인 권리에 미칠 영향에 대한 고찰

백은경
KAIST 산업디자인학과
e-mail : ekbaek@kaist.ac.kr

The impact of the EU Design Regulation Reformation on Digital Design Rights

Baek, Eunhyong
KAIST, Department of Industrial Design

요 약

2024년 3월 유럽연합(EU) 의회에서 채택되어 현재 이사회의 승인을 기다리고 있는 EU 디자인 보호규정 개정 (Regulation No 6/2002)은 디지털 디자인 권리에 중요한 영향을 미칠 것으로 예상된다. 디지털 디자인은 물리적 제품뿐 아니라 가상 환경 같은 비 물리적 환경에서도 중요한 역할을 하는 반면, 현재의 법률 체계는 이러한 영역 확장을 충분히 반영하지 못하는 문제를 가지고 있다. 본 논문은 EU 디자인 보호규정의 주요 개정 사항을 분석하고, 특히 디지털 디자인 분야에 미칠 영향을 고찰한다. 더불어, 한국에서도 빠르게 발전하는 디지털 경제에서 디지털 디자인 권리 보호를 위해 필요한 법률적 보호안의 개정을 제시한다.

The reformation of the European Union (EU) Design Regulation (Regulation No 6/2002) was adopted by the EU Parliament in March 2024 and is currently awaiting the EU Council's approval. Once implemented, the modified regularization is expected to have a significant impact on digital design rights. Digital design plays a crucial role not only in physical products but also in virtual environments, yet existing legal frameworks have struggled to adequately address this issue. This paper analyzes the key aspects of the EU Design Regulation reformation and examines its potential effects on the digital design field. Additionally, this study addresses the importance of protecting digital design rights in the rapidly evolving digital economy and proposes the need to modernize the legal framework of digital design rights protection in South Korea.

1. 서 론

2024년 3월 유럽연합 의회는 EU 디자인 보호 규정의 개정안(Regulation No 6/2002)을 승인하였다. 이 개정법의 주요 목적중 하나는 2001년에 제정되어 현재까지 쓰고있는 EU 디자인 보호 규정을 디지털 시대에 적합하도록 보완하는 데 있다. EU 디자인 보호 규정은 각국의 디자인 보호법을 존중하면서 유럽 연합 국가들을 통합한 공동 디자인 보호 규정으로 가지는 의미가 크다. 특히 이번 개정안은 그동안 디자인 산업계, 법조계에서 공통적으로 한계를 지적하고 있는 현 디자인 보호 규정을 시대에 맞게 개정하려는 첫 시도로 세계적으로 의미가 있다. 따라서 이번 개정안을 살펴보고, 이러한 변화가 디지털 디자인 보호에 영향을 미치는 부분에 대한 시사점을 고찰하는 것은 시기 적절하다고 생각된다.

2. 유럽 디자인 보호규정 개정안*

개정될 유럽 디자인 보호규정의 주요 변경 사항 중에 디지털 디자인과 관련된 부분은 다음 표와 같다:

(표 1) 디지털 디자인과 관련된 유럽 디자인 보호규정 개정(예정)

개정 대상	디자인
“디자인”의 정의 확대	“제품의 전체 또는 일부의 외관”을 포함하며, “이러한 기능의 움직임, 전환 또는 기타 모든 종류의 애니메이션”을 포함할 것.
“제품”의 정의 확대	실제 제품에 구현되지 않은 새로운 디자인의 출현과 디지털 형태로 구체화된 물체를 포함할 것.
“디자인”의 범위 확대	디자인의 범위를 “3D 프린팅”으로 확장하여 “디자인을 기록하는 모든 매체 또는 소프트웨어의 생성, 다운로드, 복사 및 제공도 디자인의 사용에 해당한다”고 명확히 함으로써 디자인의 범위를 조정할 것.

EU 디자인 규정 개정안에서 가장 주목할 부분은 디지털 디자인에 대한 보호 범위를 명확히 하고 확대했다는 점이다. 특히, "제품"과 "디자인"의 정의가 확장됨으로써, 디지털 환경에서 새롭게 등장하는 디자인의 보호가 가능해졌다. 이는 3D 프린팅, 소프트웨어, 애니메이션 등 디지털 기술을 활용한 디자인이 법적으로 보호받을 수 있는 범위가 크게 넓어졌음을 의미한다.

- **디자인의 정의 확대:** NFT 및 메타버스와 같은 기술 발전에 대응하여 "디자인" 및 "제품"의 정의가 수정될 예정이다 (Article 3의 수정). 특히, "디자인"은 디자인 기능의 움직임, 전환 또는 기타 모든 종류의 애니메이션으로 확장될 예정이다.
- **제품의 정의 확대:** 제품의 정의도 "물리적(physical) 물체로 구현되었는지 또는 비물리적(non-physical) 형태로 구체화되었는지 여부에 관계없이" 컴퓨터 프로그램 이외의 모든 산업 또는 수공예품을 포함하도록 확장될 예정이다 (Article 3의 수정). 이전 초안에서는 "디지털 형태(digital form)"라는 단어가 있었지만, 이제는 "비물리적(non-physical)" 형태까지 포함하도록 확대된 것이 굉장히 혁신적으로 판단된다. 이는 디지털 제품의 외관이 포함된다라는 점을 명확히 하기 위한 것으로 보인다. 제품을 구성할 수 있는 예시도 확장되어 매장 레이아웃과 같이 내부 또는 외부 환경을 형성하기 위한 아이템의 공간적 배치도 포함되었다. 그래픽 작품, 심블, 로고, 표면 패턴 및 그래픽 사용자 인터페이스 또한 제품의 다른 예로 추가될 예정이다.
- **보호 범위 확대:** 3D 프린팅 기술을 이용하여 복제되는 디자인 물품의 침해 문제를 보다 효과적으로 해결할 수 있도록, 디자인권의 보호 범위가 조정될 예정이다 (Article 19의 수정). 구체적으로, 디자인 제품을 복제할 목적으로 디자인을 기록하는 매체 또는 소프트웨어의 생성, 다운로드, 복사 및 제공하는 것이 디자인 권리자의 승인을 받아야 하는 디자인의 사용에 해당하는 것으로 수정됩니다.

3. 결론 및 제언

한국의 디자인 보호 체계는 현재 산업디자인진흥법을 중심으로 운영되고 있으며, 디지털 디자인 보호에 대한 명확한 규정을 두고 있지 않다. 2021년 디자인보호법 개정에서 확장디자인을 디자인 정의 규정에 포함하는 등 한국 디자인 보호법과 규정 또한 시대에 맞게 반영하려는 노력과 시도가 있지만, 우리나라의 강점인 디지털 산업과 그에 적용되는 디자인의 가치를 반영하고 법적으로 보호하기에는 아직 미흡한 부분들이 많이 있다. 특히, 디지털 제품의 설계나 3D 프린팅과 같은 새로운 기술을 고려한 보호 규정이 미비한 상황이다. 따라서, 이번 EU의 개정안은 한국의 디자인 보호 체계에 중요한 시사점을 제공한다. 한국은 디지털 경제에서 경쟁력을 유지하기 위해 이와 유사한 법적 보호 장치를 마련할 필요성이 커지고 있다.

EU가 디지털 디자인 보호 범위를 확장한 것처럼, 한국도 향후 디지털 기술의 발달에 따른 새로운 디자인 보호 요구를 수용할 수 있는 제도를 도입해야 할 시점이다. 특히, 디지털 콘텐츠, 가상 환경, 소프트웨어 디자인 등의 보호 범위를 확대하는 것이 중요할 것이다.

한국은 EU의 디자인 보호 규정 개정에서 많은 교훈을 얻을 수 있다. 특히 디지털 디자인 보호 측면에서, 한국의 법률 시스템은 기존의 물리적 제품 중심에서 디지털 환경으로 그 범위를 확장해야 한다. 이를 위해 저자는 한국의 디자인 규정 관련 기관과 법조계, 디자인 산업계에 다음 두 가지 제언을 하고자 한다:

디지털 디자인 보호 범위 확대: 3D 프린팅, 소프트웨어 디자인, 가상현실(VR) 및 증강현실(AR) 콘텐츠와 같은 디지털 디자인을 보호할 수 있는 명확한 법적 기준을 마련해야 한다.

국제적 협력 강화: 디지털 디자인 보호는 국가 간 경계를 넘어서는 문제이므로, 한국도 EU와 같은 주요 무역 파트너들과의 협력을 강화하여 글로벌 디자인 보호 체계를 구축할 필요가 있다. 이를 통해 국제적으로 통용되는 디자인 보호 기준을 수립하고, 디지털 디자인 산업의 성장을 촉진할 수 있을 것이다.

EU 디자인 보호 규정의 개정은 디지털 시대의 디자인 보호 방향을 제시한 중요한 사례이다. 한국은 이러한 변화를 면밀히 분석하고, 자국의 디자인 보호 체계를 디지털 시대에 맞춰 발전시켜 나가야 할 것이다.

*EU 디자인 보호규정 개정 (Regulation No 6/2002)안은 2024년 3월 유럽연합(EU) 의회에서 채택되어 현재 이사회의 승인을 기다리고 있으며 정확한 개정안은 학회가 개최될 무렵, 늦어도 2024년말내에 공지될 것으로 예상된다.

참고 문헌

- [1] European Commission, Proposal for a Regulation of The European Parliament and of the Council amending Council Regulation (EC) No 6/2002 on Community designs and repealing Commission Regulation (EC) No 2246/2002, November 2022
- [2] European Parliament, Revision of the Community Design Regulation in A Europe Fit for the Digital Age, Legislative Train, August 2024
- [3] EP Legislative Observatory, Industrial property: protection of Community design, 2022/0391 (COD)
- [4] European Economic and Social Committee, EESC opinion: Revision of the Design Directive and Regulation, March 2023
- [5] European Parliament, EU design protection rules ready for new technologies and circular economy, March 2024.

단안 비디오를 이용한 야구 피칭 모션 재건

김지원, 유리
아주대학교 인공지능학과
jw731@ajou.ac.kr, riyu@ajou.ac.kr

Reconstruction Baseball Pitching Motion from 2D Videos

Jiwon Kim, Ri Yu
Ajou University, Department of Artificial Intelligence

요약

우리는 2D 단안 비디오에서 야구 피칭 동작을 3D 공간에 재건하는데 발생하는 문제점을 보완하는 프레임워크를 제안한다. 이전의 연구[1]에서는 포즈 추정 기술[2],[3]을 사용하여 비디오로부터 동작을 얻어내는 방식을 사용하였기 때문에 모션 블러가 발생한 신체 부위에 대해서는 동작을 재구성하는데 어려움이 있었다. 또한 비디오에서 캐릭터의 전역 좌표를 얻을 수 없어 동작을 재건할 때 신체의 중심부가 고정된 상태로 진행되었다. 우리는 고수준의 제약 조건을 사용하여 팔에 대한 모션 블러 및 자기 폐색을 극복하였고, 신체의 전역 좌표를 추정하여 기존보다 더욱 현실성 있는 동작을 도출하였다.

1. 서론

사람의 동작 재구성 연구는 포즈 추정 기술을 통해 최근 몇 년간 놀라운 결과를 보여주고 있다. 이전 연구들은 물리 기반 시뮬레이션을 사용하여 입력 비디오에서 추정된 운동학적 포즈를 수정하였다. 그 결과, 신체의 떨림, 발 스퀘이팅, 지면 침투와 같은 문제점들이 완화되었다.

스포츠와 같은 역동적인 사람의 움직임을 촬영한 비디오에서는 플레이어가 물체를 던지거나 위치를 바꾸기 위해 갑자기 빠르게 움직인다. 이 순간, 특히 팔다리와 같은 신체의 일부가 모션 블러로 인해 흐릿하게 보이게 된다. 또한, 단안 비디오의 경우 신체의 가려짐, 즉 자기 차폐가 매우 빈번하게 발생한다. 따라서 모션 블러와 자기 차폐는 단안 비디오에서 역동적인 스포츠 동작을 재구성하는 것을 어렵게 만든다. 이러한 문제를 겪는 비디오에서 스포츠 동작에 적합한 제약 조건을 제공할 수 있다면, 타당한 동작을 생성할 수 있을 것이라고 가정하였다. 가령, 야구 경기에서 투수는 스트라이크를 얻기 위해 스트라이크 존으로 공을 던져야 하며, 이러한 제약을 우리의 프레임워크에 적용할 수 있다. 또한, 위와 같은 방법을 사용하여 전역 좌표계 안에서 신체 중심부의 적합한 움직임을 추정할 수 있다.

결론적으로 2D 단안 비디오에서 야구 피칭 동작을 물리 시뮬레이션과 심층 강화 학습을 활용하여 모션 블러와 자기 폐색 문제를 극복하고 더욱 현실성 있는 자세를 구현하였다.

2. 관련 연구

비디오에서 모션을 복원하는 것은 센서 기반 방식에 비해 효율적이다. 센서 기반 방식은 비용이 많이 들고 범위가 제한적이다. 또한 사용된 비디오의 수에 따라 접근 방식이 달

라질 수 있다. 여러 각도에서 동일한 동작을 촬영한 다수의 비디오를 활용하면 자세 복원에 큰 도움이 될 수 있지만 이 또한 다재다능하지 못하다. 따라서 단일 비디오에서 모션의 정보 제공률을 향상시킬 수 있는 방법을 사용한다.

Ri Yu[1]은 단일 비디오로부터 얻은 모션의 정보와 물리 시뮬레이션 및 심층 강화학습을 기반으로 비디오로부터 사람의 동작을 강건하게 재건하였다. 그러나 높이뛰기와 같은 역동적인 스포츠 동작의 경우 자세 추정 결과가 부정확할 때는 정확한 참조 동작을 만들어내기 위한 후보정 작업이 필요한 한계점이 존재했다.

3. 방법

우리는 Ri Yu[1]에서 스포츠 동작에 적합한 강화학습 보상을 추가하여 팔의 움직임을 강건하게 재구성하고, 전역 좌표에서 신체의 움직임을 현실성 있게 보강하였다.

3.1 공의 목표 지점 설정

우리는 투수의 비디오 영상들을 분석하였고, 공을 스트라이크 존 주변에 던지는 행위를 장려하도록 강화학습 보상을 설정하였다.

투수가 공을 던지면 곡선을 따라 움직이기 때문에 그 궤적은 예측 가능하다. 따라서 공을 던진 후 매 프레임마다 예상되는 공의 위치를 계산할 수 있다. 강화학습 내에서 공의 위치 보상 $r_{ballPos}$ 은 공이 목표 지점에 최대한 가깝게 도달하도록 즉, 캐릭터가 공을 원하는 위치에 공을 던지도록 유도하였다.

$$r_{ballPos} = \exp(-\alpha_{ballPos} \|d\|^2) \quad (1)$$

위의 식에서 d 는 공과 목표 지점까지의 거리를, $\alpha_{ballPos}$ 는

가중치이다.

투수는 스트라이크 존을 향해 공을 던져야 하므로 이를 이용, 두 벡터 \vec{p} 와 \vec{q} 사이의 차이를 최소화하는 공의 방향 보상 항목 $r_{ballDir}$ 을 고안하였다. 여기서 \vec{p} 는 공의 질량 중심 속도를 나타내는 벡터이며, \vec{q} 는 공과 목표 지점 사이의 벡터이다. $\alpha_{ballDir}$ 는 가중치이다.

$$r_{ballDir} = \exp(-\alpha_{ballDir} \|1 - \vec{p} \cdot \vec{q}\|^2) \quad (2)$$

3.2 신체의 전역 좌표 수정

Ri Yu[1]의 방식은 비디오에서 신체의 동작을 얻을 때 몸의 중심 즉, Abdomen을 고정한 상태에서 각 신체 부위의 상대적인 위치 및 속도 등을 추정하였기 때문에 신체의 이동을 측정할 수 없다. 우리는 투수가 공을 던질 때 신체의 중심이 스트라이크 존 방향으로 크게 이동하는 것을 확인하였고 공을 던지는 시점에서의 이미지에서 하체의 상대적인 위치를 계산하여 Abdomen의 전역 위치를 추정하였다.

4. 결과

특정 부위에 대한 제약 조건을 추가해줌으로써 모션 블러가 발생한 부분에 대해서도 강건하게 재건이 되어 팔을 길게 펴고 공을 원하는 위치에 던지는 모션을 그림 1, 2과 같이 시뮬레이션 할 수 있다.

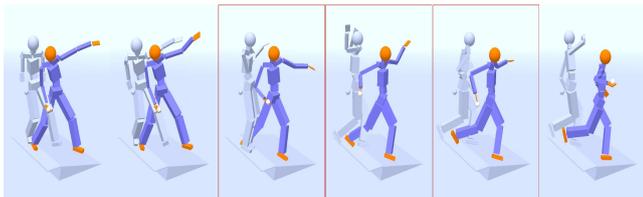


(그림 1) 시뮬레이션 결과, 측면 시점



(그림 2) 시뮬레이션 결과, 정면 시점

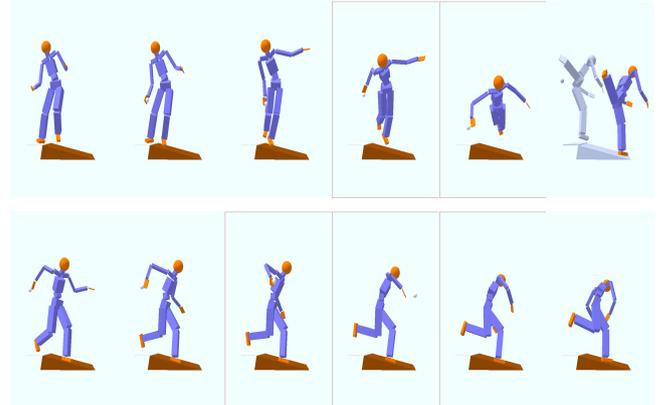
그림 3에서 팔이 제대로 추정이 되지 않아 틀린 동작이나 몸을 관통하는 등 비정상적인 자세가 있지만 이러한 문제를 해결함을 확인할 수 있다.



(그림 3) 레퍼런스 모션(회색) 및 시뮬레이션 결과 차이

위에서 언급한 프레임워크를 반영하지 않았을 때의 결과를

그림 4에서 확인할 수 있다. 첫 번째 행은 신체의 전역 좌표를 반영하지 않았을 때 캐릭터의 하체와 지면간의 충돌로 인해 균형을 잃고 쓰러져 전역 좌표를 반영했을 때와의 차이를 확인할 수 있다. 두 번째 행은 팔에 대한 강화학습 보상을 사용하지 않았을 때 모션 블러로 인해 오른팔에 대한 포즈 추정이 정확하게 되지 않아 정상적으로 팔이 뻗어 있는 문제가 발생함을 확인할 수 있다.



(그림 4) Ablation study : 첫 번째 행은 신체 전역 좌표를 사용하지 않은 결과, 두 번째 행은 팔에 대한 강화학습 보상을 반영하지 않은 결과

5. 결론

우리의 프레임워크는 모션 블러와 포즈 모호성 등의 다양한 이유로 인해 상용 포즈 추정기가 겪는 문제에도 불구하고 추정된 포즈를 바탕으로 인간의 동작을 재구성할 수 있는 잠재력을 보여준다. 비디오 안의 레퍼런스 모션의 특징을 분석하고 그에 따른 적합한 제약 조건을 이용하면 현실성 있는 모델을 구현할 수 있다.

참고문헌

[1] Ri Yu, Hwangpil Park, Jehoo Lee, Human Dynamics from Monocular Video with Dynamic Camera Movements. ACM TOG 40, 6 (2021).
 [2] CAO Z, HIDALGO MARTINEZ G, SIMON T, WEI S, SHEIKH Y.A.: Openpose: Realtime multi-person 2d pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence (2019).
 [3] KOCABAS M., ATHANASIOU N., BLACK M. J.: Vibe: Video inference for human body pose and shape estimation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (2020), pp. 5253 - 5263.

자연사 생태정보를 활용한 VR콘텐츠 연구

박기덕
유한대학교, 국립중앙과학관
want55@naver.com

A study on the VR Contents using Natural history Ecological information

park ki deok
yuhan university, national science museum

요 약

본 논문은 과학관에 보관된 전시 식물표본 채집정보의 한계점을 현장에서 채집한 신안군 희귀식물 서식지 및 생태정보를 VR영상으로 촬영하여 식물이 서식하는 위치와 생태정보를 자연사콘텐츠로 활용하여 관람자 및 연구자에게 식물생태정보의 현장감, 위치이동, 서식정보의 주변 자연 생태경관등을 활용 및 VR콘텐츠 플랫폼을 개발하여 인터랙티브한 식물생태정보의 현장감을 관람자에게 전달하고자 연구하였다. 식물 생태의 서식지와 정보를 촬영 및 자연 생태경관과 그 주변 서식지에 대한 식물 생태 정보를 획득하여 VR360의 View에서 끈끈이 귀이개의 서식 정보를 인터랙티브한 플랫폼 환경에서 2차 정보의 이미지 및 생태정보를 연출하였다. 신안군 장소를 설정하여 장소이동 및 sub target point를 지정하고 2depth 식물 생태이미지 정보로 활용하여 인터랙티브한 환경에서 식물생태환경의 주위 자연경관과 생태식물 보조정보 이미지 확대 및 관찰정보로 응용하였다. 구글맵을 연동하고 채집 및 생태정보 관찰의 이동동선을 연출하였다. 모바일 플랫폼에서 공간활용 및 Fusion studio와 Pano 2VR 프로그램을 응용하여 360이미지를 추출 및 스티칭 편집 등 인터랙션 VR콘텐츠를 제작하였다. 식물, 곤충, 새, 포유류 등 자연사 콘텐츠를 활용하여 멀티미디어 콘텐츠 및 디지털 생태정보로써 360View의 인터랙티브한 생태정보를 인식 및 활용할 수 있다.

1. 서론

1.1 연구목적

과학관에 보관된 식물전시표본의 채집정보는 관람자 및 연구자에게 식물생태정보의 생생한 현장감과 자연경관, 360기반의 주변 자연경관등을 관람할수 없는 제한적인 환경이다. 보완점으로 신안군의 끈끈이 귀이개의 희귀식물을 현장에서 VR360카메라를 활용하여 촬영하였다.

또한 신안군 자연사 콘텐츠 식물, 자연경관, 생태정보등 VR360촬영을 기반으로 자연사 생태정보를 360기반의 인터랙티브한 환경의 플랫폼을 구현하였다. 자연사 생태정보의 습득을 위한 1차원 이미지 정보의 한계를 벗어나 360View의 sub target point와 site target point 의 인터랙티브한 생태정보를 획득함으로써 다각도뷰의 자연생태정보의 실시간 관찰의 장점을 VR PC플랫폼과 VR 모바일 플랫폼으로 구현하였다. 또한 VR360 GPS좌표값과 구글map을 연동하여 모바일 환경에서 생태정보의 현재위치와 이동경로를 파악하여 Target Point로 쉽게 동선 인지 및 공간이동이 가능하다. 채집지역의 1차원 생태사진정보의 한계점을 360도 뷰의 전체 화면에서 획득할수 있는 인터랙션 생태정보 활용가능성을 제시하였다

1.2 연구범위

신안군의 채집장소로 이동하여 가상의 타겟포인트를 설정하

여 동선 및 사이트 이동, 자연경관 및 식물 생태정보를 360VR카메라를 활용하여 무선블루투스를 근거리에서 촬영 및 식물생태정보를 관찰하였다. 360카메라를 활용하여 주위 자연경관과 식물을 모니터링 하면서 1depth와 2depth의 장소이동과 부가정보에 필요한 sub target point 즉 식물이미지 확대사진을 촬영 및 습득한 서브 이미지 정보는 fusion studio 프로그램에서 스티칭 작업을 하여 파노라마 사진을 연출하고 pano2VR프로그램을 활용하여 각 장소이동과 서브 타겟포인트 작업을 한다. 각 식물정보 및 자연경관을 타겟포인트를 설정하여 동영상 정보 및 확대사진 정보를 인터랙티브한 환경으로 PC, 모바일 플랫폼으로 구현하였다.

2. 본론

2.1 끈끈이 귀이개

끈끈이 귀이개과의 여러해살이풀로 학명은 Drosera peltata 이다. 특징은 높이는 10-30cm이다. 괴경(塊莖)은 지름 6mm, 윗부분에서 가지가 갈라진다.

3. 연구 방법

3.1 제작 Process

(표 1) 제작 Process

1단계	2단계	3단계
신안군 끈끈이 귀이개 (희귀식물) 서식지 지도 Map 찾기 타겟팅 설정	신안군 끈끈이 귀이개 주변환경 VR360 촬영	VR360촬영소스 및 인터랙티브 플랫폼 제작(채집현재위 치, 끈끈이 귀이개 다각도VIEW 촬영, 각 이동동선 Site Map, 사운드, 생태정보, 서식지주변, 자연경관등)

(표 2) Plant Target Point Image

Plant information	Contents
Drosera peltata	

VR카메라에서 촬영한 전체뷰에서 끈끈이 귀이개에 근접했을시 UI(Sub target point)와 식물생태정보(G P S : 위도,경도,고도, 식물높이등)를 노출하였다[7].

(표3)VR SITE 플랫폼

Contents

360VR 영상 아래 구글Map을 연동하여 채집자의 현재위치 및 동선을 확인할수 있고, 좌측에는 각 장소이동에 대한 인터랙션 site map image 이다. VR영상 이미지에는 장소이동을 위한 Target point 링크가 있다.

(표4)VR SITE

VR Interactive Botanical Ecology Information

VR Program	Contents
Site 02 Space Move UI	
Site 03 Space Move UI	

각 인포그래픽 정보와 자연경관 생태정보, 멀티미디어 사운드 요소, 영상생태정보를 활용할수 있고, VR기기를 촬영하여 모바일 플랫폼에서 360도 실시간 현장감을 느낄수 있다.

4. 결론

과학관에 전시되고 있는 식물표본 디오라마는 채집환경의 현장감, 주변자연경관, 몰입도, 채집정보, 현장사운드, 채집위치, 채집정보를 쉽게 인지할수 없는 정보이다. 본 연구를 통해 보완점으로 신안군의 희귀식물을 현장에서360으로 촬영하여 신안군 자연사 생태정보를 VR콘텐츠로 활용하고 모바일과 PC플랫폼을 활용하여 360화면에서 현장감과 몰입도를 극대화 및 Sub image target point와 Site target point를 구현하여 생태정보 VR플랫폼을 제작하였다.

감사의 글

This research was supported by National Research Foundation of Korea (2022M3H9A1097179).

참고문헌

[1] J. H. Lee. Realistic Content (AR,VR,,360° Utilization and Improvement in Nonface to face Training Focused on Elementary School Education Journal of the The Korean Society Of Design Culture, vol. 26, no. 3, pp. 372, 2020.

[2] H. J. Lee & J. H. Choe. An Implement On Environment Responsive Smart Phone Launcher UI Using AR Camera. Journal of basic design & art, vol. 20, no. 4, pp. 425, 2019.

[3] S. C. Yoo. Tourism Marketing using VR 360 Advertising: Focusing on Content Format and Type, Novelty-Seeking Tendency, and Viewing Satisfaction. Journal of Social Science, vol. 59, no. 1, pp. 31, 2020.

은행에서 발생하는 금융 보고서 데이터의 정보 중요도에 따른 자동 등급화를 위한 ALBERT 기반 연구

전예림¹, 최예지², 이지민³, 장항배⁴
중앙대학교 일반대학원 융합보안학과^{1,2,3}
중앙대학교 경영경제대학 산업보안학과^{4*}
e-mail : {yerim1901¹, yjji9783², dlwlals³}@cau.ac.kr, hbchang@cau.ac.kr^{4*}

ALBERT-Based Study for Automatic Grading of Financial Report Data According to The Importance of Information

Jeon Ye Lim¹, Choi Ye Ji², Lee Ji Min³, Chang Hang Bae^{4*}
Department of Security Convergence, Graduate School, Chung-Ang
University^{1,2,3}
Department of Industrial Security, College of Business and Economics,
Chung-Ang University^{4*}

요 약

데이터가 경제활동의 중요한 요소로 사용되는 데이터 경제 시대에, 민감한 정보와 자금이 오가는 금융 보고서 데이터의 보호 체계 구축은 필수적이다. 이를 위해 문서의 활용도와 중요도에 따라 체계적으로 문서를 등급화하는 작업이 요구된다. 따라서, 본 논문은 불필요한 보안 인력을 줄이고 기술 유출로 인한 경제적 손실을 방지하기 위해 금융 데이터를 학습한 KB-ALBERT 기반 정보 중요도 등급화 모델을 제안한다. 본 연구는 AI-Hub의 “금융 분야 다국어 병렬 말뭉치 데이터”에서 금융 보고서만을 추출하여 사용한다. 보고서 데이터는 우선적으로 BERT 모델을 활용해 요약되며, 이후 금융 언어에 특화된 KB-ALBERT 모델을 통해 분석하여 정보의 중요도를 자동으로 등급화한다. 연구 결과, 데이터가 기밀, 비밀, 대외비, 공공의 4가지 등급으로 분류되는 것을 확인할 수 있었다. 따라서, 금융 기관은 정보의 중요도에 따라 정확하게 분류하여 보안 체계에 적용할 수 있다. 이를 통해 금융 산업에서의 보안 관리 효율성이 향상되어 경제적 손실이 방지될 것을 기대한다. 향후 연구에서는 보고서뿐만 아니라 다양한 금융 데이터를 활용하여 모델을 고도화할 예정이다.

감사의 말

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송혁신인재양성(융합보안핵심인재양성)사업의 연구 결과로 수행되었음 (IITP-2024-RS-2023-00266605)

국내 사이버 회복력 개념 정립 및 평가체계 방안 연구

박지혜¹, 이상태², 장항배^{3*}
중앙대학교 대학원 융합보안학과¹⁻²,
중앙대학교 경영경제대학 산업보안학과 교수^{3*}
e-mail : {pjh06023¹, sangtae0727², hbchang^{3*}}@cau.ac.kr

Establishing Cyber Resilience Concept and Evaluation Model in the Domestic Context

Jihye Park¹, Sangtae Lee², Hangbae Chang^{3*}
Department of Security Convergence, Graduate School,
Chung-Ang University¹⁻²
Department of Industrial Security, College of Business and Economics,
Chung-Ang University^{3*}

요 약

사이버 회복력(Cyber Resilience)은 전통적인 복원력(Resilience) 개념에서 발전하여 다양한 산업과 분야로 확산되었으며, 최근 10여 년간 급격히 변화하는 사이버 보안 환경에서 필수적인 요소로 자리 잡고 있다. 그러나 현재 국내에서는 사이버 회복력이 사이버 복원력(Cyber Recovery)과 혼용되어 지칭되는 경우가 많으며, 이를 평가할 수 있는 표준화된 체계가 미비한 실정이다. 본 연구는 사이버 회복력의 개념을 명확히 정립하고, 이를 바탕으로 국내 사이버 회복력 평가체계를 구축할 수 있는 방안을 제시하고자 한다.

본 연구는 사이버 회복력과 복원력의 개념적 차이를 구분하고, 현대 디지털 환경에서 사이버 회복력의 필요성과 중요성에 대한 심도 있는 논의를 진행하였다. 국제사회의 사이버 회복력 평가 동향을 분석하여 국내 적용가능한 실무적 방안을 모색하였으며, 심층적인 문헌 분석과 보안 전문가 자문을 통해 사이버 회복력의 개념과 평가체계를 구체화하였다. 이 과정에서 미국의 Cyber Resilience Review(CRR) 모델을 분석하여 국내에 적합한 평가지표를 개발·검증하였고, 이를 통해 조직의 대응, 복구, 적응 능력을 평가할 수 있는 지표를 도출하였다.

연구 결과, 국내 사이버 회복력 평가체계를 수립하기 위한 구체적인 방향성을 제시하였으며, 사이버 위협에 대한 조직의 대응 역량을 평가하고 향상시키는 데 기여할 수 있는 기반을 마련하였다. 이를 통해 사이버 회복력의 명확한 개념 정립과 평가 방안을 바탕으로, 국내 사이버 보안 수준을 체계적으로 평가하고 개선할 수 있는 방안을 제공할 것으로 기대한다.

감사의 말

“이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(RS-2024-00415520, 2024년 산업혁신인재성장지원사업)”

다크 패턴과 피싱: 기술 유출의 새로운 사이버 보안 위협 분석

이지현¹, 변재욱², 장항배^{3*}
중앙대학교 일반대학원 융합보안학과^{1,2}, 중앙대학교 산업보안학과^{3*}
e-mail : {ljhelloworld¹, bjw03282², hbchang^{3*}}@cau.ac.kr

Dark Patterns and Phishing: An Analysis of New Cybersecurity Threats in Technology Leakage

Ji-Hyeon Lee¹, Byun Jae Wook², Hang-Bae Chang^{3*}
Security Convergence, Graduate School, Chung-Ang University^{1,2}
Industrial Security, College of Business and Economics,
Chung-Ang University^{3*}

요약

다크패턴(Dark Patterns)은 웹사이트 및 어플리케이션에서 사용자가 충분한 정보나 선택의 여지를 제공받지 못한 상태로 의도하지 않게 비합리적 결정을 내리도록 '유도'하는 디자인적 전략을 의미한다. 이러한 수법은 피싱(Phishing)과 결합하여 새로운 사이버보안 위협을 초래할 수 있다. 피싱은 사용자를 속여 개인정보 및 민감 정보를 탈취하는 행위로, 기술 유출을 포함한 다양한 사이버 범죄에 사용되어 중대한 손실을 야기시킬 수 있다. 특히 다크패턴을 통해 강제적인 클릭을 유도하거나 가짜 경고 팝업을 통한 사용자 판단에 혼란, 불필요한 개인정보 요구 등을 실행하여 피싱 공격으로 이어질 수 있다. 본 논문은 다크패턴이 피싱 공격에 어떻게 활용될 수 있는지를 분석하고, 이 두 가지 보안 위협이 융합되어 새로운 보안 위협으로 등장할 가능성을 탐구한다. 국내에서는 다크패턴에 관한 연구가 아직 초기 단계에 있으며, 주로 전자상거래에 초점이 맞춰져 연구가 진행 되어 그 한계가 명확하다. 본 논문을 통해 다크패턴과 피싱이 기술유출 사고에서 어떻게 결합되어 새로운 사이버 보안 위협을 형성할 수 있는지 조명하고, 이에 대한 경각심을 촉구한다. 또한 이 연구는 보안 전략의 개선을 위한 기초 자료를 제공함으로써 보다 효과적인 기술 유출 대응 방안을 모색하는데 기여할 것으로 기대한다.

Acknowledgement

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구 임(RS-2024-00415520, 2024년 산업혁신인재성장지원사업)

개인정보보호 연구동향 분석: 국내 KCI 등재 학술지를 중심으로(2019-2023)

문지원¹, 박지혜², 장항배^{3*}
중앙대학교 일반대학원 융합보안학과¹⁻²
중앙대학교 경영경제대학 산업보안학과^{3*}
e-mail : {beem8728¹, pjh06023, hbchang^{3*}}@cau.ac.kr

Analysis of Research Trends in Personal Information Protection: Focused on Domestic KCI-Registered Journals

Jiwon Moon¹, Jihye Park², Hangbae Chang^{3*}
Department of Security Convergence, Graduate School,
Chung-Ang University¹⁻²
Department of Industrial Security, College of Business and Economics,
Chung-Ang University^{3*}

요 약

데이터 경제 시대에 개인정보는 보호 대상으로 간주되는 것을 넘어, 새로운 가치를 창출하는 핵심 자원으로 부상하고 있다. 이에 따라 유럽연합을 중심으로 개인정보보호 관련 법률을 제·개정하고, 인공지능과 같은 신기술 개인정보 위협에 대응하기 위한 정책을 도입하는 등 개인정보보호를 위한 국가적 차원의 노력을 증대하고 있다. 이러한 흐름 속에서 개인정보보호와 관련된 법·정책 분석, 기술적 보호 방안, 대중의 인식 변화에 관한 연구가 지속적으로 이루어지고 있으나, 최근 동향을 체계적으로 분석한 연구는 여전히 부족한 상황이다. 또한 개인정보보호는 특정 산업이나 분야에 국한되지 않고, 전반적인 비즈니스 환경과 사회에 광범위한 영향을 미치는 만큼, 다양한 분야에서 응용되는 거시적인 관점에서의 연구가 필요하다. 본 연구는 2019년부터 2023년까지 한국학술지인용색인(KCI)에 등재된 논문을 대상으로, “개인정보보호,” “개인정보 보호,” “프라이버시,” “프라이버시 보호”를 검색 키워드로 선정하여 데이터를 수집하고, 최근 5년간의 국내 개인정보보호 연구 동향을 분석하였다. 결과적으로, 본 연구는 개인정보보호에 관한 연구 흐름을 종합적으로 파악하고, 이를 기반으로 향후 개인정보보호 연구방향 설정과 학제간 연구 활성화에 기여하고자 한다.

감사의 말

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음(IITP-2024-RS-2024-00438056)

국내 공공기관 개인정보 유출 사례를 통한 PET 도입 방안 연구

이다빈¹, 배성윤², 이주노³, 장항배^{4*}
중앙대학교 일반대학원 융합보안학과¹⁻³
중앙대학교 경영경제대학 산업보안학과^{4*}
e-mail : filwoldabin97¹, baesy1994², iamjuno95³, hbchang⁴@cau.ac.kr

A Study on the Introduction of PET Based on Cases of Personal Information Leakage in Domestic Public Institutions

Dabin Lee¹, Sungyun Bae², Juno Lee³, Hangbae Chang^{4*}
Security Convergence, Graduate School, Chung-Ang University¹⁻³
Industrial Security, College of Business and Economics, Chung-Ang University^{4*}

요 약

최근 국내 대학과 법원, 관공서 등 공공기관에서 개인정보 유출 사고가 잇따라 발생하며 공공기관의 개인정보보호 중요성이 어느 때보다 부각되고 있다. 공공기관은 민감 정보를 대량으로 다루고 있어 고도화된 개인정보보호 조치가 요구된다. 본 연구는 데이터 경제 시대에서 유효한 대안으로 주목받는 PET(개인정보보호 강화기술, Privacy-Enhanced Technology) 도입을 통해 공공기관의 개인정보 보호책을 도출하고자 한다. 이를 위해 국내 공공기관의 개인정보 유출 사례를 분석하고 공공 분야에 PET를 적용한 해외 사례와 비교하여 PET 도입 방안을 모색한다. 본 연구 결과는 공공기관이 국민의 개인정보를 보호함과 동시에 안전하게 데이터를 활용할 수 있는 토대를 마련하는 데 기여할 것으로 기대한다.

감사의말

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송혁신인재양성(융합보안핵심인재양성)사업의 연구 결과로 수행되었음 (IITP-2024-RS-2023-00266605)

의사결정나무의 성능 향상을 위한 유전알고리즘 기반 이산화 방법론

류슬기, 천기정, 이철우, 최성근, 전성범
동국대학교 산업시스템공학과

e-mail : 2021112414@dongguk.edu, pridess@dongguk.edu, cjf4674@dongguk.edu,
ska07072@dongguk.edu, sbjun@dgu.ac.kr

Evolutionary Algorithm for Discretization for improving Decision Tree Performance

Seul-Gi Ryu, Gi-Jeong Chun, Chul-Woo Lee, SeongGeun Choi, Sungbum Jun
Department of Industrial and Systems Engineering, Dongguk University

요 약

제조 현장의 사물인터넷 도입으로 인해 데이터 양이 급증함에 따라, 고장 원인 분석에 기계학습 모델을 활용할 수 있는 가능성이 점점 높아지고 있다. 그러나 고장의 예측과정을 작업자가 이해할 수 없는 블랙박스 형태의 기계학습 모델은 현장에 적용하기에 다양한 한계점이 존재한다. 이러한 한계를 극복하기 위하여, 높은 예측력과 해석가능성을 지닌 기계학습 모델에 대한 연구가 필요하다. 따라서 본 연구에서는 성능을 높이기 위한 유전 알고리즘을 기반의 이산화 기법을 활용하여 전통적인 해석가능한 기계학습 모델인 의사결정나무의 성능을 개선하는 방법론을 제안한다. 제안된 방법론의 해석력과 성능을 검증하기 위하여, 다양한 데이터를 활용하여 기존의 블랙박스 알고리즘과의 성능을 비교하였다.

통합시퀀스를 활용한 시점 별 선호도 차별화 추천시스템

문상혁¹, 구자협², 이영훈³

서울과학기술대학교 데이터사이언스학과¹, 서울과학기술대학교

산업정보시스템전공², 서울과학기술대학교 산업공학과³

e-mail : noblems@ds.seoultech.ac.kr¹ rnwkguq1506@seoultech.ac.kr²,

yhoon.lee@seoultech.ac.kr³

Gated Decoupled Representation Recommendation System

Sanghyuk Moon¹, Jahyeob Koo², Younghoon Lee³

Department of Data Science, Seoul National University of Science and Technology¹,

Department of Industrial and Information Systems Engineering, Seoul National University of Science and Technology²,

Department of Industrial Engineering, Seoul National University of Science and Technology³

요약

Cross-Domain Sequential Recommendation(CDSR)은 여러 도메인의 유저 별 '아이템 상호작용 시퀀스'를 활용하여 유저의 다음 '상호작용 아이템'을 예측하는 분야이다. 기존 CDSR 방법론들은 도메인마다 개별적으로 유저의 선호도를 학습한 후, 타 도메인에 해당 정보를 전이시키는 방식으로 진행된다. 그러나 이러한 방식은 도메인 통합 아이템 시퀀스의 정보들을 학습하지 못한다는 한계를 가진다. 최근 적은 수의 연구들이 이를 극복하고자 하였으나, 도메인 통합 아이템 시퀀스 정보의 특성을 고려한 개선이 필요하다. 첫째로, 도메인 통합 시퀀스는 개별 시퀀스들과 중첩되는 정보를 가지기에, 개별 시퀀스만의 정보들을 도메인 통합 시퀀스의 정보들과 분리해줄 필요성이 있다. 둘째로, 도메인 통합 아이템 시퀀스 정보의 중요도는 시점에 따라 달라지기에, 각 시점마다 도메인 통합 시퀀스와 개별 시퀀스의 정보 반영도를 달리 가져가야 한다. 따라서 본 연구에서는 이러한 문제점을 해결하기 위해 '도메인 통합 선호도'와 '도메인 특화 선호도'를 분리하여, 도메인 통합 시퀀스에 포함된 모든 도메인의 정보를 전이시키는 동시에 개별 시퀀스에서 고유한 도메인 특성을 반영할 수 있는 'Gated Decoupled Representation Recommendation System(GDR Rec)'을 제안한다. Amazon의 Movie-Book Cross-Domain recommendation 데이터셋을 이용하여 실험이 진행되었고, 기존의 SOTA 모델을 대비 유의미한 추천 성능 지표의 향상이 있음을 확인하였다.

1. 서론

기존의 추천시스템은 주로 단일 도메인에 집중하여 책, 영화, 음악 등 여러 도메인에서의 관심사를 충분히 반영하지 못한다. 이를 해결하기 위해 CDSR이 제안되었으며, 이는 여러 도메인 간의 상호작용 시퀀스를 활용하여 더 정교한 예측을 가능하게 했다. 그러나 기존 CDSR 방법론은 도메인별로 사용자 선호도를 독립적으로 학습한 후 타 도메인에 전이하는 방식으로 진행되어, 도메인 통합 시퀀스의 정보를 효과적으로 학습하지 못하는 한계가 있다. 이를 개선하고자 CDSR with Cross-Sequence 방법론인 'C2DSR'의 경우 GNN 모듈과 attention 기법을 활용하고, 'DREAM'에서는 Transformer 인코더를 활용하여 도메인 통합 시퀀스에서 사용자 선호도를 추출한다. 그러나 두 가지 모델 모두 도메인 통합 시퀀스의 특성을 충분히 고려하지 못한다는 점에서 한계를 가진다. 예를 들어, Cross 도메인 시퀀스와 Book 도메인 시퀀스에서 동일한 아이템이 두 시퀀스에 모두 포함되어 있을 경우 개별 도메인의 고유한 선호도를 정확히 반영하지 못한다. 또한, 도메인 통합 시퀀스의 정보 중요도는 시점에 따라 달라질 수 있

는데 기존 방법론은 이러한 시간적 변화에 적절히 대응하지 못한다. 따라서 본 연구에서는 이런 문제를 해결하기 위해 '도메인 통합 선호도'와 '도메인 특화 선호도'를 분리하여, 도메인 통합 시퀀스의 모든 도메인의 정보를 전이시키면서 동시에 개별 시퀀스의 고유한 도메인 특성을 반영할 수 있는 'Gated Decoupled Representation Recommendation System (GDR Rec)'을 제안한다.

2. 관련 연구

2.1 CDSR

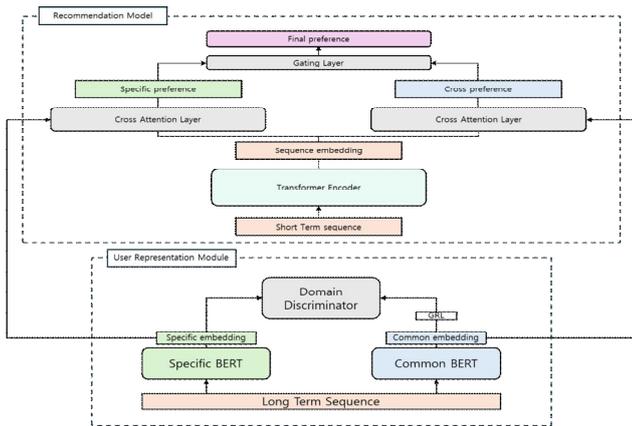
Sequential Recommendation(SR)은 사용자의 과거 상호작용 시퀀스를 기반으로 미래의 행동을 예측하는 방법론이다. Cross-Domain(CD)은 한 도메인에서 학습된 사용자 선호도를 다른 도메인으로 전이하여, 새로운 도메인에서의 성능 저하 혹은 데이터 부족 문제를 해결하는 기술이다. 따라서 CDSR은 이 두 가지를 결합하여, 여러 도메인의 시퀀스 데이터를 통합적으로 분석하여 보다 정확한 추천을 가능하게 한다.

2.2 CDSR with Cross Sequence

1)'C2DSR'(Contrastive Cross-Domain Sequential Recommendation)은 대조학습을 적용하여 도메인 간의 상호작용을 균형있게 반영하고, 각 도메인 상호작용을 고려하여 시퀀스를 생성한 모델이다.

2) 'DREAM'은 Extraction Attention Module을 활용하여 단일 도메인과 교차 도메인 사용자 선호를 동시에 학습하는 접근 방식을 제시한다. 또한, Supervised Contrastive Learning과 Focal Loss 기법을 도입해 데이터 희소성과 클래스 불균형 문제를 해결한 모델이다.

3. 방법론



(그림 1) 전체 아키텍처 구조

3.1 User Representation Module

그림 1의 이 모듈에서는 Specific BERT와 Common BERT는 사용자 상호작용의 Long Term Sequence를 입력으로 받아 MLM을 수행한다. 이를 통해 각각의 시퀀스 정보가 반영된 Specific Embedding과 Common Embedding이 생성된다. 이후 Domain Discriminator는 앞에서 생성된 두 Embedding을 Label에 따라 분류한다. 핵심은 Specific Embedding은 해당 label에 따라 명확하게 구분되도록 학습되고, Common Embedding은 도메인 간 구분이 어렵게 학습된다. 이 때 Common BERT는 역전파시 Gradient Reverse Layer(GRL)를 통해 -loss를 받아 결론적으로 특정 도메인에 특화된 정보와 도메인 간 공통된 정보가 서로 명확하게 구분되며 학습된다.

3.2 Recommendation Module

추천 모듈에서는 사용자의 Short Term Sequence를 입력으로 받아 Transformer Encoder를 통해 임베딩 벡터로 변환한 후, Cross Attention Layer를 이용해 시퀀스 임베딩과 Specific/Common Embedding 간의 교차 연산을 수행하여 각각 Specific Preference 벡터와 Common Preference 벡터를 생성한다. Gating Layer는 각 시점별로 정보 반영도를 학습하여 이 두 벡터 중 추천에 더 중요한 정보를 판별하여 Final Preference를 도출한다. 최종적으로 이 선호도는 모든 아이템과의 유사도를 계산하여 예측 아이템의 확률값을 도출하게 된다.

4. 실험결과

4.1 데이터셋

포괄적이고 공정한 실험을 위해 널리 사용되는 Amazon의 Movie-Book dataset을 사용한다. 또한 본 연구에서는 C2DSR 논문의 실험설정을 따른다.

4.2 실험설정

공정한 비교를 위해, 10개 미만의 상호작용을 가진 항목과 단일 도메인에서만 상호작용한 사용자들도 제외한다. 또한 CDSR의 순차적 제약을 충족시키기 위해 각 도메인 내에서 세 개 미만의 항목을 가진 시퀀스도 제외한다.

Negative sampling과 유사한 방식을 이용하여 총 1000개의 후보군을 구성했다. 이는 실제로 맞춰야 하는 target값 1개와 완전히 랜덤하게 샘플링된 데이터 999개로 이뤄져 있다. 하이퍼파라미터 설정은 두 모듈에 대해 다르게 적용시켰다. User Representation Module에서는 1)시퀀스 길이 : 100, 2)학습률 : 0.001, 3)배치 사이즈 : 4, 4)epochs : 100, 5)alpha : 0.5로 설정했다. 여기서 5)alpha는 domain loss와 MLM loss간의 조정 가중치이다. 나머지 다른 하나인 Recommendation Module에서는 1)시퀀스 길이 : 10, 2)학습률 : 0.001, 3)배치 사이즈 : 2048, 4)epochs : 100으로 설정하였다.

4.3 실험결과

		movie	
		NDCG@10	HR@10
CDCF	NCF-MLP	2.96	5.3
	CoNet	3.01	5.35
SDS	SASRec	3.69	5.2
	GRU4Rec	3.73	5.4
	SR-GNN	3.78	5.81
CL	CL4Rec	6.2	10.31
	CoSeRec	6.42	11.12
CDS	Pi-Net	4.17	6.11
	PSJNet	4.76	7.53
	C2DSR	6.09	9.91
	DREAM	<u>11.11</u>	<u>18.79</u>
Our	GDR Rec	11.942	21.318

(표 1) 실험결과

표 1를 살펴보면, 기존 SOTA 모델인 DREAM 모델의 NDCG@10은 11.11이고, HIT@10은 18.79이지만, 우리 모델의 성능 결과는 NDCG@10은 11.942이고, HR@10은 21.318로 유의미한 성능 향상을 확인할 수 있다.

참고문헌

[1] Jiangxia Cao, Xin Cong, Jiawei Sheng, Tingwen Liu, and Bin Wang. 2022. Contrastive Cross-Domain Sequential Recommendation.

[2] Xiaoxin Ye, Yun Li, and Lina Yao. 2023. DREAM: Decoupled Representation via Extraction Attention Module and Supervised Contrastive Learning for CrossDomain Sequential Recommender.

소규모 및 대규모 고객 집합을 가진 VRPPD 문제 해결을 위한 Greedy 알고리즘 및 하이브리드 유전자 알고리즘의 비교 분석

차병현, 이현경, 김해중*
경기대학교 산업시스템공학과

e-mail : malong0313@naver.com, hk0325lee@naver.com,
haejoong.kim@kyonggi.ac.kr

Comparative Analysis of Greedy Search, and Hybrid Genetic Algorithm for Solving VRPPD with Small to Large Customer Sets

Byeonghyun Cha, hyunkyung Lee, Haejoong Kim*
Department of Industrial Systems Engineering, Kyonggi University

요 약

VRP(Vehicle Routing Problem)은 물류 및 배송, 배달 서비스 등에서 운영 효율을 높이기 위해 자주 연구되고 있는 분야이다. 본 논문에서는 VRP의 변형 중 하나인 CVRP (Capacitated VRP)와 VRPPD (VRP with Pickup and Delivery)을 혼합한 개념의 모델에서의 최적화를 진행했다. Greedy Search, Hybrid Genetic Algorithm 등 2가지 방법론 적용을 통해 수요별 각 방법론의 성능과 계산 소요 시간등을 비교 분석하였다.

1. 서론

VRPPD (Vehicle Routing Problem with Pickup and delivery)는 픽업과 배송이 고려된 차량 경로 할당 문제로 여러 고객에게 물품을 전달하기 위해 최소 비용으로 차량의 경로를 최적화하는 VRP (Vehicle Routing Problem)의 변형이다. VRP는 물류 및 배송, 배달 서비스 등에서 효율성을 높이기 위해 자주 연구되고 있는 분야이다. 최근 전 세계적으로 탄소 배출 감소 능력과 ESG (Environment, Social, Governance) 경영이 기업의 지속 가능성을 평가하는 중요한 요소로 일반적으로 경로의 길이, 차량 용량, 시간 제약 등의 다양한 제약 조건을 고려하여 실제 운영 환경과 유사한 모델을 구축한다. VRPPD와 같이 제약 조건이 늘어날수록 문제가 복잡해지고 이는 NP-hard 문제가 된다.[1]

2. 관련 연구

픽업 및 딜리버리 문제(VRPPD)의 다양한 변형은 물류 및 운송 최적화 분야에서 중요한 연구 주제로 다루어져 왔다. 각 변형은 문제의 구조에 따라 다양한 알고리즘을 통해 최적화가 이루어졌으며, 수익 극대화, 비용 최소화, 시간 창 고려 등의 목표를 달성하고자 하는 노력이 반영되었다.

2.1 1대1 VRPPD(1-1-VRPPD)

Li and Chen (2016)의 연구는 캐리어 협력 시 일부 요청을 선택적으로 서비스함으로써 수익을 극대화하는 시나

리오를 연구하였다.[2] Chami et al.(2019)의 연구는 시간 창과 페어링된 수요를 고려한 선택적 픽업 및 딜리버리 문제를 다루며, 수익 최대화와 비용 최소화라는 두 가지 목표를 동시에 고려한 최적화 기법을 제안했다.[3]

2.2 다대다 VRPPD (M-M-VRPPD)

Hernández-Pérez and Salazar-González (2004)는 다중 상품의 픽업 및 딜리버리 문제를 해결하기 위해 Branch-and-Cut 알고리즘을 사용하였으며[4], Zhao et al. (2009)는 유전자 알고리즘을 적용하여 다중 상품 픽업 및 딜리버리 문제를 해결하고, 우수한 수렴 성능을 보였다.[5]

2.3 1대다-다대1 VRPPD (1-M-1-VRPPD)

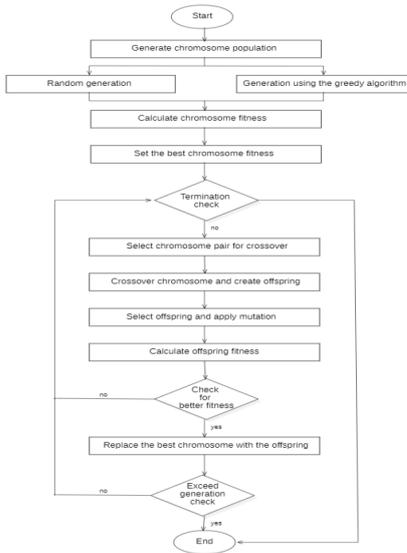
Gribkovskaia et al. (2008)는 선택적 픽업을 포함한 딜리버리 문제(Single Vehicle Routing Problem with Deliveries and Selective Pickups)를 연구하였으며, 개선된 시뮬레이티드 어닐링 알고리즘을 사용하였다.[6] Assis and Maravilha (2013)는 다목적 반복적 지역 탐색 알고리즘을 적용하여 선택적 픽업과 고정된 딜리버리를 최적화하였고[7], Coelho et al. (2015)는 변수 이웃 탐색(VNS) 알고리즘을 통해 선택적 픽업 및 딜리버리 문제를 해결하였다.[8]

이와 같이, 다양한 VRPPD 변형 문제에 대한 연구는 각기 다른 알고리즘과 최적화 기법을 통해 문제를 해결하고 있으며, 본 연구는 이러한 선행 연구들을 바탕으로 고객 노드 수를 기준으로 소,중,대규모 상황에서 VRPPD 문제의 최적화를 목표로 한다.

3. 실험

3.1 Genetic Algorithm

다음은 유전알고리즘 적용 순서이다.



(그림 2) Genetic Algorithm Flow Chart

유전 알고리즘은 먼저 초기 개체군을 설정하는 단계로 시작된다. 이 단계에서는 인구(population) 크기, 세대(generation) 수, 돌연변이 확률, 차량 용량 등의 값이 정의되며, 이는 알고리즘의 실행 조건을 결정짓는 중요한 요소이다. 본 논문에서는 유전 알고리즘과 하이브리드 유전 알고리즘 두가지의 유전 알고리즘을 사용했다. 두 알고리즘의 차이는 초기 개체군 생성 방식에 있다.

3.2.1 초기 개체군 생성 (Generate Chromosome Population)

유전 알고리즘의 초기 개체군은 전부 무작위로 생성되는 반면, 하이브리드 알고리즘의 초기 개체군은 하이브리드 방식을 통해 생성된다. 개체군의 절반은 무작위로 생성되며, 나머지 절반은 greedy 알고리즘을 사용해 경로를 형성한다. 이 과정에서 greedy 알고리즘은 각 단계에서 가장 가까운 위치를 선택하여 경로를 결정함으로써 초기 개체군을 보다 효율적으로 구성한다.

3.2.2 적합도 계산 (Calculate Chromosome Fitness)

생성된 각 개체의 적합도를 계산하는 단계에서는 경로의 총 거리를 고려하며, 차량의 용량 제약을 반영한다. 만약 경로 상의 수요량이 차량의 최대 용량을 초과하게 되면, 차량은 허브로 복귀한 뒤 새로운 경로를 시작해야 한다. 이처럼 차량 용량을 고려한 총 이동 거리로 각 경로의 효율성이 평가되며, 이를 바탕으로 적합도가 결정된다.

3.2.3 부모 선택 (Select chromosome pair for crossover)

부모 개체는 적합도 점수를 바탕으로 선택된다. 적합도가 높을수록 선택될 확률이 높아지며, 이 과정에서 룰렛휠 선택 방법이 사용된다.

3.2.4 교차 (Crossover)

선택된 부모 개체 간에는 순서 교차(ordered crossover)가 이루어져 자손 개체가 생성된다. 자손은 부모의 유전

정보를 일부 물려받아 새로운 경로를 구성하게 된다.

3.2.5 돌연변이 (Mutation)

이후 교차가 완료된 자손에게는 돌연변이가 적용된다. 돌연변이 과정에서는 스왑 돌연변이(swap mutation)를 통해 경로의 일부가 무작위로 변경된다. 이는 경로의 두 위치를 교환하는 방식으로 이루어진다.

3.2.6 자식의 적합도 계산 (Calculate offspring fitness)

새롭게 생성된 자손의 적합도를 계산한다.

3.2.7 더 나은 적합도 확인 (Check for better fitness)

자손이 기존의 최고 개체보다 더 나은 적합도를 보이는지 확인한다.

3.2.8 최고 염색체를 자식 염색체로 교체 (Replace the best chromosome with the offspring)

만약 자손이 부모보다 더 좋은 적합도를 가진다면, 최고 개체는 자손으로 대체된다.

3.2.9 세대 초과 확인 (Exceed Generation Check)

알고리즘이 설정된 최대 세대 수를 초과하지 않았다면, 부모 선택, 교차, 돌연변이의 과정을 반복하여 새로운 세대를 생성한다.

3.2.10 반복 (Repeat)

종료 조건을 충족하지 않았다면, 부모 선택, 교차, 돌연변이 과정을 반복하면서 새로운 세대를 생성합니다.

3.2.11 종료 (End)

알고리즘이 종료 조건을 충족하면, 최종적으로 발견된 가장 적합한 경로(최적 경로)가 출력된다. 즉 차량 용량 제약을 고려하여 총 이동 거리가 최소화된 최적의 경로가 최종 출력된다.

3.3 실험 데이터

본 논문에서는 실험을 위한 데이터를 수요의 정도를 달리하며 각 방법론들을 적용하였다. 수요의 정도는 고객 노드의 수를 2배수로 조정하며 small, medium, large 고객 집단으로 설정하였다. 차량의 대수는 모든 수요에서 2대로 통일 하였고, 두 차량의 종류는 같은 종류(용량)의 차량(Homogeneous fleet)으로 설정하였다.

참고문헌

[1] Berbeglia, G., Cordeau, JF., Gribkovskaia, I. et al. Static pickup and delivery problems: a classification scheme and survey. TOP 15, 1 - 31 (2007).
 [2] Yuan Li, Haoxun Chen, Christian Prins, Adaptive large neighborhood search for the pickup and delivery problem with time windows, profits, and reserved requests, European Journal of Operational Research, Volume 252, Issue 1, 2016, Pages 27-38,
 [3] Al Chami Z, Manier H, Manier M-A. A lexicographic approach for the bi-objective selective pickup and delivery problem with time windows and paired demands[J]. ANNALS OF OPERATIONS RESEARCH, Dordrecht: Springer, 2019, 273(1 - 2): 237 -

255.

[4] Hernández-Pérez H, Salazar-González J-J. A branch-and-cut algorithm for a traveling salesman problem with pickup

and delivery[J]. Discrete Applied Mathematics, 2004, 145(1): 126 - 139.

[5] Zhao F, Li S, Sun J, 等 . Genetic algorithm for the one-commodity pickup-and-delivery traveling salesman problem[J]. Computers & Industrial Engineering, 2009, 56(4): 1642 - 1648.

[6] Gribkovskaia I, Laporte G, Shyshou A. The single vehicle routing problem with deliveries and selective pickups[J].

Computers & Operations Research, 2008, 35(9): 2908 - 2924.

[7] Assis L P, Maravilha A L, Vivas A, 等. Multiobjective vehicle routing problem with fixed delivery and optional

collections[J]. OPTIMIZATION LETTERS, Heidelberg: Springer Heidelberg, 2013, 7(7): 1419 - 1431.

[8] Coelho I M, Munhoz P L A, Haddad M N, 等. A hybrid heuristic based on General Variable Neighborhood Search for the Single Vehicle Routing Problem with Deliveries and Selective Pickups[J].

Electronic Notes in Discrete Mathematics, 2012, 39: 99 - 106.

AI 자율 제조 환경에서의 SDF(Software Defined Factory) 구축 전략에 관한 연구

하태원, 최서운, 강현민, 황규선

울산대학교 산업경영공학부

e-mail : xodnjs7965@mail.ulsan.ac.kr, cseoyun@mail.ulsan.ac.kr,
kanghm0124@mail.ulsan.ac.kr, gysunhwang@ulsan.ac.kr

A Study on the Strategy for Building SDF (Software Defined Factory) in an AI Autonomous Manufacturing Environment

Taewon Ha, Seoyun Choi, Hyunmin Kang, Gysun Hwang
School of Industrial Engineering, University of Ulsan

요 약

최근 자율 주행 차량의 수요가 증대됨에 따라 자동차 내에 다양한 소프트웨어 기능들이 추가된 SDV(Software Defined Vehicle) 개념의 차량이 개발되고 있다. 또한 국내 완성차 제조 공장에서는 모든 생산 요소를 데이터로 연결하고, 목적 기반형 자동차(PBV, Purpose Built Vehicle)를 적시에 생산하는 제조 운영 체계를 만들기 위한 다양한 시도를 하고 있다. 우리나라 정부와 자동차 업계에서는 이러한 공장의 형태를 SDF(Software Defined Factory)라는 개념으로 정의하고, 관련 기술 발굴, 시스템 설계 등에 대한 전략을 수립 중에 있다. 소수의 논문과 정책 연구에서 SDF 개념에 대한 정의와 현황을 기술하였으나, SDF 개념에 대한 정확한 정의나 목적성을 정리한 연구는 미비한 실정이다. 이에 본 연구는 소프트웨어 중심의 자동차를 의미하는 SDV를 생산하는 고도화된 스마트공장 개념인 SDF의 개념, 목적성, 필요 기술 등을 정의하고자 한다. 본 연구를 통해 자동차 산업의 새로운 공장의 형태인 SDF에 대한 개념 정립 및 방향성에 대한 기초 자료로 활용될 수 있을 것으로 기대한다.

자율운반로봇의 관리 고도화를 위한 정체 관리 KPI 발굴

황규선

울산대학교 산업경영산업안전공학부

e-mail : gyusunhwang@ulsan.ac.kr

Identification of congestion management KPI for Automated Guided Vehicle

Gyusun hwang

School of industrial engineering, University of Ulsan

요 약

현재 제조 공장에서는 물류 자동화 기술 수준이 높아짐에 따라 기존의 작업자 기반 물류 운반 대신 자율 운반 로봇(AGV, Automated Guided Vehicle)을 활용한 물류 방식이 점차 확대 적용되고 있다. AGV는 레일 기반 운행 방식으로 정해진 레일을 통해 이동하는 방식으로 이를 관리하는 시스템인 ACS(Agv Control System)은 최적화된 물류를 관리하기 위해 개발되었다. ACS는 실시간 물류 이동 관리, 경로 최적화 등의 기능을 수행하고 있으며, 특히 물류 현황 관리를 위해 물류 이동 시간, AGV 부하율 등의 KPI를 선정하여 관리하고 있다.

본 연구에서는 정해진 레일을 통해 이동하는 AGV 운행 방식의 특성을 고려하여 정체가 발생할 경우에 대한 심각성 수준을 인지하기 위한 정체 관리 KPI를 발굴하는 것을 목표로 한다. 이를 위해 정체 발생 시뮬레이션을 개발하고, 정체 관리KPI를 적용하여 정체 수준을 파악하는 KPI에 대한 설계 및 검증을 진행한다. 또한 제안하는 KPI의 검증을 하기 위하여 물류 지표들과의 상관성을 분석하여 정체관리 지표의 효과성을 검증하여 실용성있는 정체 KPI를 제안하고자 한다.

병렬 챔버를 가진 반도체 클러스터 장비의 lot 투입 스케줄링 문제

황성보¹, 노윤서¹, 이가영¹, 하가영¹, 한준희^{1*}

¹부산대학교 산업공학과

e-mail : ssbb7977@pusan.ac.kr, 1204yunseo@naver.com, lgy011123@naver.com,
hgy2124@naver.com, junhan@pusan.ac.kr

A Scheduling problem for semiconductor cluster tools with parallel chambers

Seongbo-Hwang, Yunseo-Noh, Gayoung-Lee, Gayoung-Ha, Junhee-Han
Department of Industrial Engineering, Pusan National University

요 약

본 연구는 lot 분할, 허용 가능한 챔버 및 자원 제약을 고려하여 makespan과 lot opening time을 최소화하는 반도체 챔버 장비의 스케줄링 문제를 다룬다. 연구 대상은 반도체 8대 공정 중 하나인 식각 공정에서 사용되는 병렬 챔버를 가진 클러스터 장비이다. 각 챔버는 parallel machine으로 볼 수 있으며, lot은 허용 가능한 챔버에 분할되어 처리될 수 있다. 또한 설비의 load port의 개수에 따라 클러스터 장비 내에서 동시에 작업 될 수 있는 lot의 개수가 제한된다. 이 문제를 해결하기 위해 혼합정수선형계획법을 제시하며 현장에서 요구하는 제한된 시간 안에 문제를 효과적으로 풀기 위하여 modified simulated annealing (MSA) 알고리즘을 제시한다. 제시한 알고리즘의 성능은 다양한 크기의 실험에서 혼합정수선형계획법, 디스패칭 룰과 비교를 통해 평가하였다. 평가 결과, 제시한 알고리즘이 디스패칭 룰 대비 좋은 성능을 내는 것을 확인하였다. 본 연구 결과는 반도체 식각 공정 및 제품 포장 공정의 생산계획 수립에 적용될 수 있을 것으로 기대된다.

스마트폰 지자기 센서를 활용한 마그네틱 액세서리 식별 연구

최재민, 김효수
중앙대학교 컴퓨터공학과
e-mail : jaeminld@cau.ac.kr, hskimhello@cau.ac.kr

Magnetic accessory identification using an off-the-shelf smartphone's magnetometer.

Jaemin Choi, Hyosu Kim
School of Computer Science and Engineering, Chung-Ang University

요 약

Magsafe가 iPhone에 탑재되면서, 기기의 후면에 자석 기반으로 마그네틱 지갑, 충전기, 거치기, 스마트폰 쿨러 등 다양한 액세서리를 부착하는 것이 가능하게 되었다. 이 논문에서는 이러한 물리적인 마그네틱 액세서리들과 스마트폰을 소프트웨어적으로 연결하는 MagID 시스템을 소개한다. 예를 들어 마그네틱 액세서리가 스마트폰에 부착된다면, 스마트폰에서 각 액세서리마다 지정해 놓은 특정 어플리케이션을 실행시키거나 특정 일을 수행하도록 할 수 있다. 이러한 기능적인 연결을 가능하게 하는 MagID의 핵심 아이디어는 스마트폰에 어떠한 마그네틱 액세서리가 부착 혹은 탈착 되었는지를 식별하는 것이다. 구체적으로 말하자면, 우리는 각 마그네틱 액세서리가 각기 다른 자기적 성질을 가지고 있는 것을 관측했고, 이에 따라서 자기장을 센서로 관측함으로써 마그네틱 액세서리를 식별하고자 했다. 하지만 지자기 센서만을 이용한다면, 사용자의 움직임으로 인한 노이즈와 환경이 변함에 의해 생긴 주변 자기장 변화가 영향을 줄 수 있다. 따라서 MagID에 비자기적인 특성으로 가속도 센서, 자이로 센서와 충전 상태를 이용하여 노이즈를 제거하고 MagID의 견고성 및 성능을 높이고자 했다. 연구의 실험에서 MagID에 영향을 줄 수 있는 환경에서 다양한 변수를 고려한 실험을 진행했고, 그 결과 MagSafe가 탑재된 다양한 스마트폰과 노이즈가 있는 실험 환경에서 높은 성능을 보여주는 것을 확인했다.

웨어러블 인터페이스를 이용한 RF 신호 기반 접촉 부위 추정 기법 연구

신수현, 김효수
중앙대학교 컴퓨터공학과
e-mail : girinssh@cau.ac.kr, hskimhello@cau.ac.kr

RF signal-based contact estimation techniques for wearable interfaces

Suhyeon Shin, Hyosu Kim
School of Computer Science and Engineering, Chung-Ang university

요 약

가상현실과 증강현실 기술의 발전하고 상용화됨에 따라 다양한 기기들이 출시되고 많은 사용자들에게 보급되었다. 이러한 기기들은 사용자의 자세 인식과 다양한 입력을 위해 카메라를 사용함과 동시에 손으로 잡고 사용하는 인터페이스 기기를 사용한다. 하지만 이 인터페이스 기기는 카메라로 가려지는 부분은 인식할 수 없고, 양 손이 자유롭지 못하다는 단점이 있다. 또한 가상현실 기기로 인해 현실의 공간을 인지하기 힘든 상황에서 사고가 났을 때 대처하지 못할 수 있다.

이러한 문제를 해결하기 위해 본 논문에서는 처프 신호 형태의 전자파를 송수신하는 팔찌형 인터페이스 기기를 개발하고 이 신호를 처리하여 접촉에 사용된 손과 그 부위를 추정하는 기법을 제안한다. 이 인터페이스 기기는 샘플링 후에도 주파수 대역이 겹치지 않도록 설계된 서로 다른 처프 신호를 내보내고, 이 신호를 수집해 주파수 성분을 추출한다. 이 주파수 성분을 서포트 벡터 머신의 피쳐로 사용하여 접촉 부위와 접촉한 손을 추정한다. 5밀리초의 처프 신호를 사용했을 때, 10개의 부위에 대해서 부위 당 40개의 처프 신호, 총 400개의 데이터만으로 훈련시켰음에도 불구하고 부위 당 1036개의 데이터를 검증했을 때 정확도 약 98%를 달성하였다.

혼합 현실 기기를 활용한 현실과 가상 공간 간의 좌표 매핑 기술 연구

성락빈¹, 이대원^{1*}

¹중앙대학교 예술공학대학 응용예술공학과

*교신저자

e-mail : rakbin007@naver.com, dwlee@cau.ac.kr

A study on coordinate mapping technique between real and virtual spaces using mixed reality device

Rakbin Sung¹, Daewon Lee^{1*}

¹Department of Applied Art and Technology, College of Art and Technology, Chung-Ang University

*To whom correspondence should be addressed

요 약

본 연구는 혼합 현실 기기와 클라우드 컴퓨팅 서비스를 활용하여 현실과 가상 공간 간의 좌표를 매핑하고 이를 바탕으로 혼합 현실 기기로부터 다른 기기 내의 가상 공간으로의 확장을 위한 프로세스를 제안한다. 홀로렌즈를 사용하여 혼합 현실 환경을 구축하고 Azure Spatial Anchor 서비스를 이용하여 가상 공간 내의 앵커 좌표를 현실 공간에 고정하도록 하는 좌표 매핑 프로세스를 제안하며, 홀로렌즈 이외에 기기를 통해 새로운 가상 공간을 생성하고 로컬 서버 통신과 좌표 매핑 프로세스를 통해 홀로렌즈 내의 가상 공간을 타 기기의 가상 공간으로 확장할 수 있도록 하는 프로세스를 제안한다. 본 연구에서 제안한 가상 공간 확장 프로세스를 활용하여 혼합 현실 기기의 장점을 극대화하고 다양한 산업 분야에서의 혼합 현실 기기 사용 방안을 제시할 수 있을 것으로 기대한다.

1. 서 론

혼합 현실(Mixed reality) 기기는 다수의 카메라를 비롯한 고성능의 센서를 통해 현실 세계를 인식하고 가상 세계와 현실 세계 간 상호작용이 가능하게 하는 공간 컴퓨팅(Spatial computing) 기술을 통해 사용자에게 몰입감 있는 경험을 제공한다.

혼합 현실 기기의 성능이 발전함에 따라 보다 정밀한 공간 컴퓨팅 기술을 구사할 수 있게 되었으며, 다양한 분야에서 이를 활용할 수 있는 연구가 진행되고 있다 [1]. 또한, 혼합 현실 기기에 클라우드 컴퓨팅 기술을 접목하여 더 정밀하고 빠른 공간 컴퓨팅 기술이 제안되고 있다 [2].

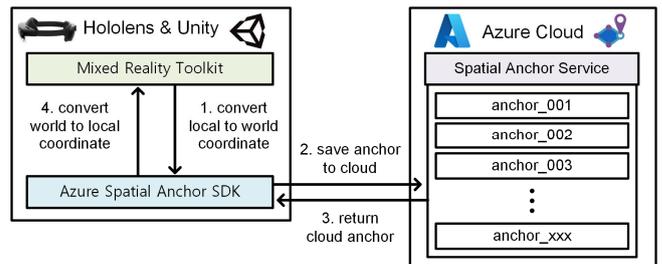
본 연구는 혼합 현실 기기와 클라우드 컴퓨팅 서비스를 기반으로 하여 공간 컴퓨팅을 위한 현실과 가상 공간 간 좌표 매핑 기술을 분석 및 개발한다. 또한, 혼합 현실 기기와 타 컴퓨팅 기기와의 통신을 통해 여러 가상 공간을 하나의 좌표계로 통합하는 공간 확장 기술을 제안한다.

2. 본 론

본 연구에는 혼합 현실 기기로 Microsoft 社의 Hololens(이하 홀로렌즈)를 사용한다. 또한, Mixed Reality 라이브러리와 3D 게임엔진인 Unity(이하 유니티)를 활용하여 홀로렌즈를 위한 MR 프로젝트를 개발한다. 가상 공간과 현실의 좌표를 공유하기 위해 Microsoft 社의 클라우드 컴퓨팅 서비스인 Microsoft Azure를 사용한다.

Azure의 Spatial anchor는 현실 공간의 좌표를 클라우드를 통해 공유하여 가상 공간 생성 시 클라우드의 좌표를 참조할 수 있도록 한다.

2.1 좌표 매핑 프로세스

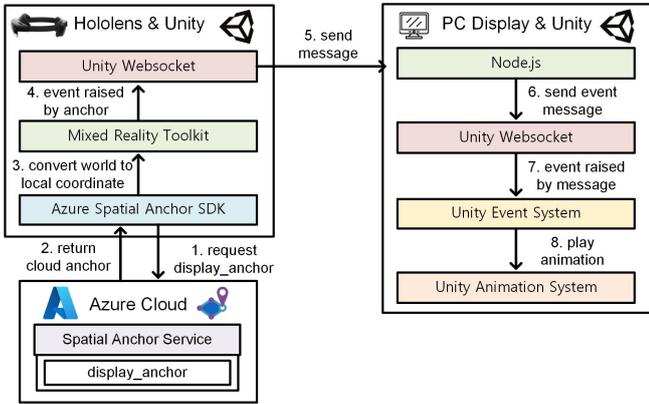


(그림 1) 좌표 매핑 프로세스

그림 1은 홀로렌즈와 공간 앵커 클라우드 서비스를 활용하여 현실과 가상 공간 간 좌표를 매핑하는 과정을 나타낸 것이다. 유니티를 통해 홀로렌즈 센서 활용을 위한 Mixed Reality Toolkit과 클라우드 서비스 사용을 위한 Azure SDK를 포함한 3D MR 프로젝트를 생성한다. 홀로렌즈를 통해 프로젝트의 로컬 좌표에 앵커를 등록할 수 있다. Mixed Reality Toolkit은 로컬 좌표로 등록된 앵커를 월드 좌표로 변환하여 Azure SDK에 전달한다. Azure 클라우드 서비스에 등록된 사용자명과 토큰을 통해 클라우드에 앵커의 좌표를 등록할 수 있다. 반대로 클라우드에

등록된 앵커의 좌표를 획득할 수도 있다. Azure SDK를 통해 획득한 앵커는 게임엔진 레벨 내의 지역 좌표로 변환되어 홀로렌즈 상에 표현된다. 본 프로세스를 통해 홀로렌즈 착용자는 현실의 고정된 위치를 항상 참조할 수 있고 가상 오브젝트를 현실 좌표에 보다 정확하게 매핑할 수 있다.

2.2 가상 공간 확장 프로세스



(그림 2) 좌표 매핑 기반 공간 확장 프로세스

그림 2는 좌표 매핑과 타 기기와의 통신 기술을 활용하여 다른 가상 공간으로 공간을 확장하는 프로세스를 나타낸 것이다. 본 프로세스에는 다른 가상 공간을 생성할 수 있는 별도의 디바이스를 필요로 하며 두 디바이스에는 같은 에셋을 공유하는 게임엔진 프로젝트가 동작하고 있는 것으로 가정한다. 그림 2는 홀로렌즈 이외에도 유니티를 통해 가상 공간을 구현한 PC와 이를 표현하는 디스플레이를 가정하였다. Azure의 공간 앵커 서비스를 이용하면 디스플레이의 현실 좌표를 클라우드에 업로드 해놓을 수 있다. 한 번 업로드된 좌표는 디스플레이를 이동시키지 않는 한 고정되어 홀로렌즈에서 이를 활용할 수 있다. 홀로렌즈에서 Azure Spatial Anchor SDK를 통해 Azure Cloud에 접속하고 업로드된 디스플레이의 좌표를 획득한다. Mixed Reality Toolkit을 통해 월드 좌표를 홀로렌즈 게임엔진 레벨 내의 로컬 좌표로 변환한다. 이상의 프로세스를 통해 디스플레이의 현실 좌표를 홀로렌즈의 가상 좌표에 표현한다. 디스플레이를 통해 표현되고 있는 PC의 유니티 레벨과 홀로렌즈의 레벨 간 통신을 위해서 두 기기가 모두 접속할 수 있는 로컬 서버를 구축하였다. Node.js를 활용하여 PC에 로컬 서버를 구축하고 유니티의 Websocket 패키지를 통해 홀로렌즈와 PC 두 기기 모두 로컬 서버에 접속하도록 하였다. 홀로렌즈의 유니티 레벨에서 디스플레이의 앵커 위치를 참조하여 이벤트를 발생시키고 이를 로컬 서버를 통해 PC의 유니티 레벨로 메시지를 전송할 수 있다. 두 기기의 프로젝트가 동일한 에셋을 공유하고 있다고 가정하므로 메시지 통신을 통해 홀로렌즈 에셋의 애니메이션과 PC의 에셋 애니메이션 타이밍을 조절하여 두 가상 공간이 서로 연결된 것 같은 효과를 만들어낼 수 있다.



(그림 3) 공간 확장 프로세스 예시

그림 3은 공간 확장 프로세스를 사용하여 홀로렌즈의 레벨에서 PC의 레벨로의 공간 확장을 구현한 예시이다. 그림3에서 홀로렌즈 레벨의 에셋이 디스플레이로 이동하고 있으며 (좌상단) 디스플레이를 뚫고 내부로 진입하는 모습을 보인다 (우상단). 이 때 홀로렌즈 내에서는 좌표 매핑을 통한 디스플레이 앵커로부터 이벤트가 발생하고 해당 이벤트는 로컬 서버를 통해 PC에 메시지를 전송한다. 그림 3의 좌측 하단 그림에서 에셋이 디스플레이 공간 내부로 진입한 것을 확인할 수 있다. 디스플레이에 메시지가 전송되어 이벤트가 발생하고 미리 준비된 동일한 에셋의 애니메이션이 재생되어 홀로렌즈 내의 에셋이 PC의 가상 공간으로 이동한 것처럼 보인다. 그림3의 마지막 그림에서 에셋이 완전히 디스플레이 내 공간으로 이동해 남은 애니메이션을 재생하는 것을 확인할 수 있다.

3. 결 론

본 연구에서 제안한 프로세스가 다양한 분야에서 활용될 수 있을 것으로 기대한다. 제조 산업에서 현실의 제조 공정과 가상 공간의 좌표를 매핑하여 작업 지원 기기로서 혼합 현실 기기의 정밀도를 상승시킬 수 있을 것이다. 또한, 광고 또는 콘텐츠 분야에서 가상 공간 확장 프로세스를 적절히 활용한다면 기존의 일반 디스플레이나 VR, AR 등의 기기에서는 보여줄 수 없는 뛰어난 몰입감을 소비자에게 보여줄 수 있을 것으로 기대한다.

참 고 문 헌

[1] Khoshelham, Kourosh, Ha Tran, and Debaditya Acharya. "Indoor mapping eyewear: geometric evaluation of spatial mapping capability of HoloLens." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 42, 2019. pp.805-810.
 [2] Delmerico, Jeffrey, et al. "Spatial computing and intuitive interaction: Bringing mixed reality and robotics together." IEEE Robotics & Automation Magazine 29.1, 2022. pp.45-57

음성 기반 상호작용을 위한 디지털 휴먼 시스템 연구

우성미¹, 이대원^{1*}

¹중앙대학교 예술공학대학 응용예술공학과

*교신저자

e-mail : warm10216@cau.ac.kr, dwlee@cau.ac.kr

A study on digital human systems for speech-based interaction

Seongmi Woo¹, Daewon Lee^{1*}

¹Department of Applied Art and Technology, College of Art and Technology,
Chung-Ang University

*To whom correspondence should be addressed

요 약

최근 몇 년 간 정보 통신 기술의 급속한 발전은 디지털 휴먼 기술의 진보를 크게 촉진하였다. 고성능 컴퓨팅, 인공지능(AI), 자연어 처리(NLP), 음성 인식 및 컴퓨터 비전 등 다양한 기술이 발전하면서 디지털 휴먼은 더욱 정교하고 현실적으로 진화하고 있다. 이러한 기술은 사용자에게 개인화된 경험을 제공할 수 있도록 하고, 인간과의 상호작용에서 보다 자연스럽게 몰입감 있는 경험을 제공할 수 있도록 한다. 특히, 실시간 음성 인식과 감정 분석 기술이 결합되면서 디지털 휴먼은 사용자와의 맞춤형 상호작용을 통해 교육, 의료, 서비스 등의 분야에서 사용자 만족도를 높이는 데 중요한 역할을 하고 있다. 본 연구에서는 이러한 발전을 기반으로 디지털 휴먼 제작에서 사용되는 음성 기반 상호작용 시스템을 개발하고 구현하는 데 중점을 둔다. 최종적으로 본 연구는 게임 엔진을 활용한 디지털 휴먼 환경에서 고도화된 언어 모델 및 최신 자연어 처리 기술을 활용하여 사용자 맞춤형 상호작용을 제공하는 시스템을 설계하였다.

1. 서론

4차 산업혁명 이후 정보 통신 기술이 발전하며 디지털 트윈 기술 또한 주목받고 있다. 특히 디지털 트윈 기술은 인간컴퓨터 상호작용(Human-Computer Interaction, HCI) 분야에서 큰 가능성을 제시하고 있다.[1] 디지털 휴먼은 이러한 디지털 트윈 개념을 인간의 가상 표현에 적용한 것으로, 가상 환경 내에서 인간과 유사한 상호작용을 수행할 수 있는 가상의 인격체를 의미한다. 디지털 휴먼은 사용자와의 상호작용에서 감정, 언어, 동작 등을 이해하고, 이를 기반으로 실제와 유사한 반응을 제공함으로써 보다 자연스럽게 몰입감 있는 경험을 가능하게 한다. 최근 기술의 발전으로 몰입감 있는 경험 뿐 아니라 사용자에게 개인화된 경험을 제공하는 디지털 휴먼의 연구 또한 활발하게 진행되고 있다.[2]

음성 기반 상호작용은 디지털 휴먼 시스템의 핵심 요소 중 하나로, 사용자가 음성 명령을 통해 디지털 휴먼과 직접적으로 소통할 수 있다. 최근 AI 기술의 발전으로 음성 인식의 정확도와 처리 속도가 크게 향상되면서, 디지털 휴먼과의 음성 기반 상호작용이 다양한 산업 분야에서 실질적인 응용 가능성을 갖추게 되었다. 예를 들어, 디지털 휴먼은 교육, 의료, 고객 서비스와 같은 분야에서 사용자 맞춤형 상호작용을 제공하며, 이를 통해 학습 효과를 극대화하거나 고객 경험을 개선하는 데 활용될 수 있다. 본 연구

는 디지털 휴먼 시스템에서 음성 기반 상호 작용 시스템을 개발하는 것에 집중한다.

2. 본론

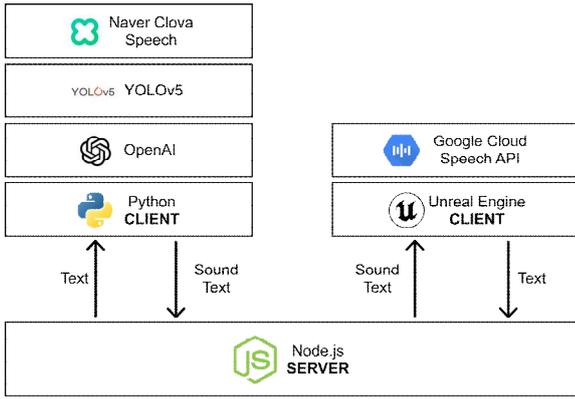
2.1 디지털 휴먼 환경 구축

본 연구는 사용자와 상호작용을 진행할 디지털 휴먼을 제작하는 것에서 시작된다. 이를 구현하기 위해 에픽 게임즈(Epic Games)에서 개발한 게임 엔진인 언리얼 엔진(Unreal Engine)의 Meta Human Creator을 사용하였다. 언리얼 엔진을 기반으로 디지털 휴먼의 외형을 제작하였기 때문에 디지털 휴먼과의 상호작용 시스템 또한 언리얼 엔진을 기반으로 제작하였다.[3]

2.2 음성 기반 상호작용을 위한 시스템 구축

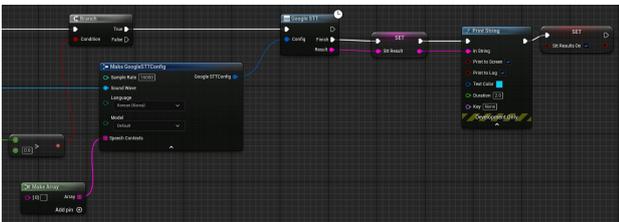
언리얼 엔진의 디지털 휴먼과 현실 사용자간의 음성 기반 상호작용을 위해 그림 1과 같은 시스템을 설계하였다. 음성 기반 상호작용 시스템은 Node.js를 통해 구현한 서버, 파이썬으로 구현된 클라이언트 및 언리얼 엔진에서 구현된 클라이언트로 구분된다. Node.js 서버는 로컬 환경에서 실행되는 서버로, 언리얼 엔진과 파이썬 코드 사이의 연결을 활성화하기 위해 사용된다. Node.js의 socket.io를 사용하여 통신을 활성화한다.

다음은 언리얼 엔진의 클라이언트이다. 언리얼 엔진의 클라이언트는 Google Cloud Speech To Text(STT)를 활용하여 사용자의 음성을 텍스트로 변환한 후 Node.js의



(그림 1) 음성 기반 상호작용 시스템 설계도

서버로 전송한다. STT 기능은 다음과 같이 블루프린트를 이용하여 구현하였다. 그림 2는 언리얼 엔진에서 블루프린트를 통해 구현된 STT 기능을 보여준다. 이후 STT를 통해 얻은 텍스트 데이터를 Node.js 서버로 전송하여 파이썬 클라이언트에서 활용할 수 있도록 한다.



(그림 2) 언리얼 엔진에서의 STT 구현 블루프린트

파이썬 클라이언트에서는 OpenAI의 GPT, Yolov5 및 Naver Clova Speech 등을 활용하여 입력받은 사용자의 텍스트를 바탕으로 디지털 휴먼의 답변을 생성한다. GPT 프롬프팅을 통해 디지털 휴먼의 인격을 구축한다. 파이썬 서버로 전달된 사용자의 텍스트는 GPT API에 전달되고, 최종적으로 디지털 휴먼의 대답 텍스트로 변환된다. 해당 텍스트는 Naver Clova Speech API 를 통해 음성으로 변환되며, 해당 음성은 Node.js 서버를 통해 언리얼 엔진 클라이언트로 전달된다. 파이썬 클라이언트에서 사용되는 Yolov5 모델은 웹캠을 통해 사용자의 성별 및 특징을 인식하여 GPT 답변 생성에 사용될 수 있도록 한다. 이를 통해 개인화 된 사용자 경험을 줄 수 있도록 시스템을 구축하였다.

최종적으로 얻어진 음성 데이터는 Node.js 서버를 통해 언리얼 엔진으로 전달된다. 언리얼 엔진은 해당 음성을 입력으로 하여 디지털 휴먼의 입모양을 생성하고 입모양 생성이 완료되었다면 음성과 함께 입모양 애니메이션을 재생한다. 그림 3은 대화를 하고 있는 디지털 휴먼의 모습을 보여준다.

4. 활용 가능성

본 연구에서 개발한 음성 기반 디지털 휴먼 상호작용 시스템은 여러 분야에서 실질적인 응용 가능성을 제공하며, 사용자 맞춤형 경험을 극대화하는 데 기여할 수 있다.

의료 분야에서 환자와 상호작용을 통해 정서적 지원을 제공하거나 적절한 피드백을 제공함으로써 환자에게 보다 자연스러운 상담 경험을 제공할 수 있다. 엔터테인먼트 분야에도 활용되어 더욱 실감 나는 사용자 경험을 제공할 수도 있다. 또한 고객 서비스 분야에도 활용되어 사용자와 24시간 실시간으로 소통할 수 있는 챗봇의 역할을 수행할 수 있다. 기존 챗봇과 달리 사용자의 감정과 의도를 파악하여 더 자연스러운 대화를 제공할 수 있고, 개인화된 서비스를 제공하여 사용자에게 보다 나은 경험을 제공할 것이라 예상된다.



(그림 3) 대화를 하고 있는 디지털 휴먼

5. 결론

본 연구에서는 디지털 휴먼과 사용자의 상호작용 시나리오를 가정한 후, 음성 인식 및 자연어 처리 기술을 활용하여 사용자 맞춤형 경험을 제공하는 시스템을 구현하였다. 특히 음성만을 입력으로 하는 것이 아니라 사용자의 웹캠 이미지를 입력 데이터로 하여 보다 세세한 개인화 경험을 줄 수 있도록 시스템을 설계하였다. 이러한 연구 결과는 교육, 의료, 엔터테인먼트 등 다양한 분야에서 디지털 휴먼 시스템의 응용 가능성을 탐색하고, 향후 디지털 휴먼 기술의 발전 방향을 제시하는 데 기여할 수 있을 것이다.

참고문헌

[1] WANG, Baicun, et al. Human Digital Twin in the context of Industry 5.0. Robotics and Computer-Integrated Manufacturing, 2024, 85: 102626.
 [2] THITE, Mohan. Digital human resource development: where are we? Where should we go and how do we go there?. Human Resource Development International, 2022, 25.1: pp.87-103.
 [3] FANG, Zhixin; CAI, Libai; WANG, Gang. MetaHuman Creator The starting point of the metaverse. In: 2021 International Symposium on Computer Technology and Information Science (ISCTIS). IEEE, 2021. p. 154-157.

이벤트 데이터 기반 인터랙티브 스포츠 스트리밍 웹 서비스 개발

신동민¹, 이대원¹

¹중앙대학교 예술공학대학 응용예술공학과
email : min1125@cau.ac.kr, dwlee@cau.ac.kr

Development of an Interactive Sports Streaming Web Service Based on Event Data

Dongmin Shin¹, Daewon Lee¹

¹Department of Applied Art and Technology, College of Art and Technology, Chung-Ang University

요약

기존 스포츠 중계 서비스는 일방적인 정보 제공에 그쳐 시청자와의 상호작용이 부족한 한계가 있다. 이를 해결하기 위해 본 연구에서는 객체 인식 기술과 이벤트 데이터를 활용하여 경기 중 발생하는 다양한 상황을 시각화하고, 사용자가 직관적으로 접근할 수 있는 인터랙티브 요소를 제공하는 새로운 스트리밍 서비스를 제안한다. 제안된 서비스는 데이터 중심의 UI/UX 설계를 통해 사용자에게 실시간 경기 통계 정보, 선수 위치 및 정보, 이벤트 시각화 등을 제공하여 축구 중계의 몰입감과 정보 접근성을 극대화한다. 본 연구는 이러한 데이터 기반의 새로운 스포츠 스트리밍 서비스의 가능성을 제시하며, 인터랙티브 웹 스트리밍 서비스를 위한 소프트웨어 설계 방법을 제안한다.

1. 서론

최근 스포츠 중계 서비스는 일방적인 정보 제공 방식을 넘어, 데이터를 기반으로 한 다양한 시각화와 정보를 제공하는 새로운 접근이 요구되고 있다. 특히 축구와 같은 대중적인 스포츠에서는 시청자의 관심과 몰입도를 유지하는 것이 중요한데, 기존의 스포츠 중계는 실시간으로 시청자의 요구를 충족시키지 못하는 한계가 있다. 전통적인 스포츠 중계는 해설자 중심의 정보 전달에 의존하며, 시청자의 배경 지식에 따라 정보 전달의 수준이 달라지는 문제가 있다. 이에 따라 일반적인 시청자에게는 이해하기 어려운 설명이 제공되거나, 반대로 기존 스포츠 팬들에게는 충분히 심층적인 분석이 제공되지 않는다는 문제점이 발생한다.

본 연구는 이러한 문제를 해결하기 위해 이벤트 데이터 기반의 인터랙티브 축구 스트리밍 웹 서비스를 개발하여 시청자의 축구 시청 경험을 개선하는 것을 목표로 한다. 이를 위해 선수 위치 데이터를 수집하는 객체 인식 기술과 축구 경기 중 발생하는 상황을 메타데이터로 기록한 이벤트 데이터를 결합하여 실시간 데이터를 구축하였다. 제공되는 서비스는 해당 데이터를 기반으로 경기 중 발생하는 상황들을 시각화하고, 사용자가 시청 중 상호작용을 통해 정보를 얻을 수 있는 다양한 기능적 요소를 지원한다.

2. 관련 연구

2.1 AI 기반 스포츠 분석

스포츠 경기에서 인공지능(AI)을 활용한 데이터 분석은 경기 성과를 평가하고 선수의 퍼포먼스를 측정하는 데 중

요한 역할을 한다. AI를 활용한 분석은 경기 전략 수립, 선수의 능력 평가, 전술 분석 등에서 매우 유용하며, 다양한 연구들이 이를 통한 성과를 입증해왔다[1-2]. 이러한 AI 기반 분석 방법들은 스포츠 경기 데이터를 보다 효과적으로 활용하여 경기 결과를 예측하거나 선수의 향후 성과를 평가하는 데 유용하다[3-4]. 본 연구는 기존 연구들이 AI를 활용한 경기력 예측과 전략적 결정에 중점을 둔 것과는 달리, AI 기반 객체 인식과 실시간 이벤트 데이터 시각화를 결합하여 변화하는 경기 상황을 시각화하는 방향으로 접근하고 있다.

2.2 스포츠 데이터 시각화와 스트리밍 서비스

스포츠 분야에서는 데이터 시각화와 사용자 친화적 디자인을 통해 시청자의 몰입감을 높이고, 사용자의 니즈를 충족시키기 연구가 이루어지고 있다[5-6]. 데이터 시각화는 경기 정보를 직관적으로 이해할 수 있도록 돕고, 시청자의 경기 이해도를 높이는 데 중요한 요소로 작용한다. 특히, 실시간 경기 데이터의 시각적 표현은 사용자의 참여도를 높이고, 더 나은 스포츠 중계 경험을 제공하는 데 필수적이다. 이에 따라, 본 연구에서는 데이터 시각화와 인터랙티브 사용자 경험 설계를 통해 스트리밍 서비스의 사용자 몰입도를 향상시킬 수 있는 방안을 제시하고자 한다.

3. 연구 방법 및 결과

3.1 웹 UI/UX 디자인

본 연구의 웹 UI/UX 디자인은 경기 이벤트 데이터를 시각화하고 각 컴포넌트를 통해 정보를 제공함으로써, 사용자가 인터랙티브하게 정보를 확인할 수 있도록 설계되

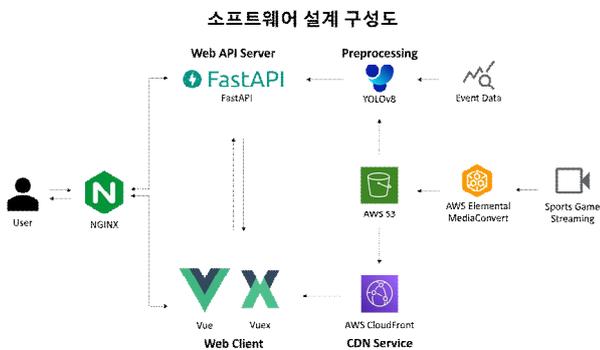
었다. 웹 화면은 그림 1과 같이 구성되어 있다. 각 선수를 따라다니는 총 4개의 시각화 패널로 구성되어 있으며, 이를 통해 선수 능력치 비교, 실시간 기록, 공격 시퀀스, 슈팅 데이터 등의 정보를 시각적으로 제공한다. 또한, 보로노이 시각화와 패스 성공률 예측을 영상 위에 구현하여 사용자가 원하는 영상을 선택해서 시청 가능하도록 제공한다. 하단에는 선수들의 상세 정보를 확인할 수 있는 컴포넌트가 배치되어 있다.



(그림 1) 축구 스트리밍 웹 서비스 화면

3.2 웹 소프트웨어 설계

전체적인 소프트웨어 구조는 그림 2와 같이 웹 서비스 운영과 동시에 데이터 가공과 스트리밍이 가능하도록 설계하였다. 웹 어플리케이션은 프론트엔드로 Vue.js를 백엔드로 FastAPI를 채택하여 구현하였다. 스트리밍 기능은 AWS의 CDN 서비스를 활용하였고 선수 위치 데이터 수집을 위해 객체 인식 모델로 YOLOv8을 활용하였다.



(그림 2) 축구 스트리밍 웹 서비스 설계 구성도

3.3 데이터 파이프라인

데이터 파이프라인은 그림 2와 같이 실시간 축구 경기 스트리밍, 이벤트 데이터 수집 상황을 가정하고 설계하였다. 스트리밍 영상은 AWS Elemental MediaConvert를 통해 트랜스코딩되어 S3에 저장되며, AWS CloudFront을 이용한 CDN을 통해 사용자에게 제공된다. S3에 저장된 영상을 기반으로 YOLOv8 모델을 통해 객체 인식 및 라벨링 작업이 실시간으로 수행되며, 수집된 이벤트 데이터와 같이 Web API 서버로 전송된다. 최종적으로 가공된

데이터는 시각화되어 사용자에게 제공되며, 보로노이 다이어그램과 같은 시각화 기법을 활용하여 경기 내 선수들의 움직임과 전략적 정보를 제공한다. 이러한 데이터 파이프라인은 실시간 데이터의 처리와 전달을 스트리밍과 동시에 가능하도록 설계되었으며, 데이터 기반 스포츠 스트리밍 서비스를 구현하는 방법을 제시한다.

4. 결론

본 연구는 축구 경기의 실시간 스트리밍 상황과 이벤트 데이터 수집 상황을 가정하여 인터랙티브 스포츠 스트리밍 웹 서비스를 구현하였다. 전통적인 스포츠 중계의 사용자 경험을 개선하기 위해 이벤트 데이터 기반 UI/UX 디자인 및 기능 구현을 시도하였으며, 실시간 스트리밍을 위한 웹 소프트웨어 설계, 데이터 파이프라인을 구축하여 새로운 방식의 스포츠 스트리밍 웹 서비스 실현 가능성을 제시하였다. 향후 연구에서는 다양한 종목에서의 적용 가능성을 검토하고, 실제 사용자 피드백을 수집하여 서비스 개선 가능성을 탐구하고자 한다.

참고문헌

[1] Cossich, V. R., Carlgren, D., Holash, R. J., Katz, L. "Technological breakthroughs in sport: Current practice and future potential of artificial intelligence, virtual reality, augmented reality, and modern data visualization in performance analysis." *_Applied Sciences_*, Vol.13, No.23, 2023, pp.12965.

[2] Tuyls, K., Omidshafiei, S., Muller, P., Wang, Z., Connor, J., Hennes, D., Hassabis, D. "Game Plan: What AI can do for Football, and What Football can do for AI." *_Journal of Artificial Intelligence Research_*, Vol.71, 2021, pp.41-88.

[3] 송화봉, 박영식, 홍재원, 왕설. "축구선수 경기력예측 시스템 구축을 위한 유효 변수 추출에 관한 연구: 중국 프로축구리그를 중심으로." *_한국정보기술학회논문지_*, 제21권, 제7호, 2023, pp.115-122.

[4] 김민중, 이현아. "경기 기록 통계와 뉴스 기사에 대한 자연어처리를 결합한 야구 승부 예측 시스템." *_디지털콘텐츠학회논문지_*, 제24권, 제5호, 2023, pp.1041-1047.

[5] Kokkinakis, A. V., Demediuk, S., Nölle, I., Olarewaju, O., Patra, S., Robertson, J., Block, F. "Dax: Data-driven audience experiences in esports." *_Proceedings of the 2020 ACM International Conference on Interactive Media Experiences_*, 2020, pp.94-105.

[6] 서효민, 김상아. "프로스포츠 구단의 ICT 융합 기술에 대한 소비자 태도 연구: 확장된 기술수용모델(ETAM)을 중심으로." *_디지털콘텐츠학회논문지_*, 제21권, 제3호, 2020, pp.521-528.

AI를 이용한 게이미피케이션 교수학습법 연구 모형 설계

윤봉식¹, 김진경²

남부대학교 자동차기계공학과¹, 목포대학교 아트앤디자인학부²

e-mail : jscomi@nambu.ac.kr¹, 0326jinkyung@daum.net²

Design of a Research Model for Gamification Teaching and Learning Methods Using AI

Bong-Shik Yun¹, Jin Kyung Kim²

Department of Automotive & Mechanical Engineering, Nambu University¹

Department of ArtnDesign, Mokpo National University²

요 약

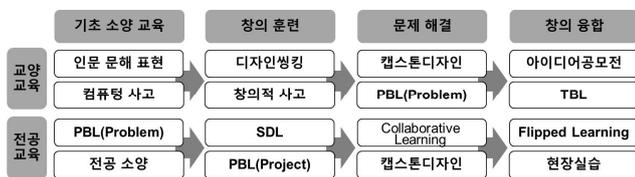
In the AI era, creative and integrative education is increasingly recognized as a critical component for not only the efficiency of knowledge cultivation but also for the active development of learners' job skills. Accordingly, this study has been developing a research model from 2022 to the present, focusing on gamification teaching and learning methods that apply AI, specifically tailored for Undergraduate students who are familiar with devices and social environments.

The application of AI technology in teaching and learning is approached from the perspectives of both learners and instructors. Based on this distinction, the study established five hypotheses that set the research direction and objectives. Through a multi-faceted research collaboration process, including literature reviews and expert discussions, a five-year research model encompassing both sample studies and applied research was developed.

Additionally, by presenting a research performance management model, the study aims to contribute to enhancing the quality management of teaching and learning in higher education. It also seeks to provide integrated research recommendations, ultimately aiming to improve students' learning outcomes and conduct empirical research on AI-applied teaching methods for instructors.

1. 서 론

생성형 AI의 출현 이후 대학교육에서도 AI시대에 적합한 창의융합적 교수학습법의 변화가 요구되고 있다. 현재 국내 대학에서 운영 중인 창의융합관련 교과는 창의, 문제해결, 창의융합 팀학습을 포함하는 구조로 구성되어 있다. 관련 교과들은 기술트렌드 수용성이 높으나 일부 교과들은 매뉴얼 방식을 선호하고 있어서 각 과정들의 유기적 연계를 고려한 교수학습법 변화가 절실하다.



(그림 1) 교양 및 전공 내 창의융합 교육체계

그림 1과 같이 연구 대학의 전공과 교양교과 내 편제되어 있는 창의융합형 교과와 편성을 살펴보면 교수자들의 소통 연계나 학습자의 경험을 적층해 나가기 쉽지 않은 구조를 가지고 있다. 전공 교과에서 팀기반 및 프로젝트기반 학습의 실무 효과검증은 가능하지만 교양 교과에서 학습자 전공 다양성이나 성향 차이, 융합의 문제들로 프로젝트기반 또는 팀학습 적용에 어려움이 있는 것으로 조사되었다.

다.[1] 또한 코로나시대를 지나며 다자 간의 협의나 토론 학습 문화가 2019년 이전에 비해 현저히 떨어지고 있어 실무 산업에 배치되는 2022~2024년 졸업생들의 협업능력과 소통능력에 부정적 피드백이 높아지고 있고, 학습자들의 학습방법과 교수법에 대한 만족도 역시 일방적인 제시형 교육보다는 쌍방향 교육에 대해 보다 불편함을 느끼는 것으로 조사되었다. 이에 학생들의 사용자경험환경에 적합한 학습과정 내 즐거움과 세대 친화적 AI기술의 교수학습법 적용 연구가 기획되었다.

2. 관련 연구

2.1 수요자 경험환경 변화

현대 대학생들은 디바이스와 플랫폼UX에 익숙하고 게임과 SNS를 통한 소비활동과 소통에 편안함을 느끼며 게이미피케이션에 친숙하게 수렴한다. 선도기업들의 AI 콘텐츠 내 게이미피케이션 요소는 학습 효익이 높아서 이미 대학들의 AI 활용 교과 개설과 AI플랫폼을 이용한 교육기반 구축이 선제적으로 이루어지고 있는 추세이다.[2] MIT 테크놀로지리뷰가 최근 2년 선정한 10대 미래기술은 AI기술과 연계해 분석, 생성, 어시스턴스, 설계영역에 밀접하고 이는 산업요구인 근미래 취업 역량으로 이해된다.

중국 크로스보더 전자상거래 플랫폼이 소비자의 구매의도에 미치는 영향 연구 -중국 온라인 플랫폼B2C형을 중심으로-

제 1 저자 : JIN TIANLE (진티엔러)
전남대학교 대학원
디지털미래융합서비스협동과정
e-mail : 2239522687@qq.com
교신저자 : 박복재
전남대학교 글로벌비즈니스학부
e-mail : bjpark73@jnu.ac.kr 박복재

A Study on the Impact of China's Cross Border E-Commerce Platform on Consumers' Purchase Intentions - Based on the B2C type of online platform in China-

JIN TIANLE Bok-Jae, Park
Chonnam National University
Digital Future Convergence Service Cooperative Course

요 약

최근 몇 년 동안 크로스보더 전자 상거래는 중국의 지속적인 발전과 성장을 동시에 중국 대외 무역의 발전을 촉진하는 새로운 방법이 되었다. 크로스보더 전자 상거래의 발전을 장려하고 지원하기 위해 정부는 크로스보더 전자 상거래 정책을 지속적으로 수립하고 조정하여 크로스보더 전자 상거래의 자유로운 발전을 위한 편안한 정책 환경을 조성함과 동시에 국경 간 전자 상거래의 표준화되고 질서 있는 발전을 보장하여 더 많은 소비자가 국제 표준에 부합하도록 한다. 크로스보더 전자 상거래 플랫폼의 어떤 요소가 소비자의 구매 의향에 영향을 미치는지 연구하는 것은 크로스보더 전자 상거래 기업이 그에 따라 플랫폼을 개선하고 크로스보더 전자 상거래 기업의 발전을 촉진하며 중국 크로스보더 전자 상거래의 발전을 촉진하여 중국 대외 무역의 발전을 촉진하는 데 도움이 다. 이 기사는 먼저 크로스보더 전자 상거래의 발전 상황, 크로스보더 전자 상거래 플랫폼의 유형 및 국경 간 전자 상거래의 문제에 대해 설명한다. 구매의사의 정의와 그 영향요인에 대한 국내외 학자들의 연구문헌을 정리하여 이성행위이론, 계획행위이론, 기술수용이론 세 가지 소비자행동이론 모델을 분석하여 본 논문의 연구 과제에 적합한 기초모형을 기술수용이론으로 선정하였으며, 기술수용모형에 플랫폼 이미지, 플랫폼 품질, 쇼핑 경험 등 세 가지 외적요인의 연구기반은 위의 외적요인이 유용감지, 사용 용이감지, 위험감지의 중개작용을 통해 소비자의 쇼핑의사에 영향을 미치는 모델을 구축하였으며, 이를 바탕으로 모델의 변수를 가정하였다.

1. 서 론

최근 몇 년 동안 정보 기술의 발전과 주민들의 생활 수준의 향상으로 크로스보더 전자 상거래가 빠르게 발전했다. 중국 크로스보더 전자상거래 시장 규모는 2021년에 처음으로 1조 달러를 초과하여 1.2조 달러에 도달했고, 2023년까지 중국 크로스보더 전자상거래 시장 규모는 1.4조 달러에 달하여 2019년에 비해 약 2배 증가하였으며, 연평균 복합 성장률은 18.5%로 같은 기간 중국 GDP의 연평균 복합 성장률 6.2%보다 높았다. 2028년까지 중국 크로스보더 전자상거래 시장은 연평균 복합 성장률 11.3%로 안정적인 성장하여 2.4조 달러에 이를 것으로 예상된다. 중국의

크로스보더 전자 상거래가 발전함에 따라 소비자는 더 많은 경로를 통해 외국 상품을 구매할 수 있으며, 출장이나 해외여행, 대형 쇼핑몰, 해외 쇼핑 사이트 해외 직구, 지인 대리 구매 채널 외에도 소비자들은 점차 국경 간 전자 상거래 플랫폼을 이해하고 익숙해지고 있다. 동시에 우리나라의 다양한 국경 간 쇼핑 전자상거래 플랫폼이 우후죽순처럼 등장하고 있으며, 국내 크로스보더 전자상거래 대기업들은 티몰 글로벌(Global), 징동(JD) 글로벌(Global) 쇼핑 등 해외 구매와 글로벌 구매 사업을 잇달아 선보여 더 많은 고품질의 해외 상품을 소비자에게 제공할 수 있다. (아이리서치 연구원,2024).

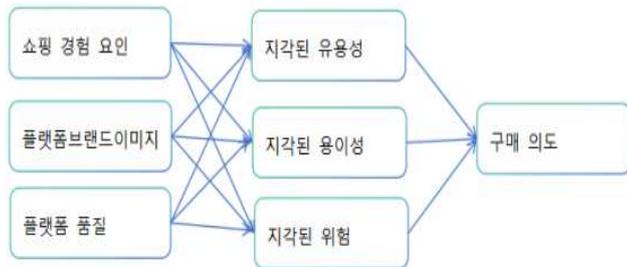
또한 젊은이들의 소비 잠재력이 지속적으로 향상됨에 따라 크로스보더 전자 상거래 산업의 급속한 발전을 위한 새로운 추진력을 제공했다. 젊은 소비자들은 촉달, 구매, 충성도가 높은 사람들 사이에서 높은 성장률을 보였고 구매 전환 효율도 높았다. 크로스보더 전자상거래 웹사이트의 급속한 발전으로 사용자의 특성을 이해하고 자체 네트워크 자원을 사용하여 크로스보더 전자상거래의 콘텐츠를 개선한다. 이러한 방식으로 크로스보더 전자 상거래 쇼핑 의향의 주요 원인에 대한 연구가 특히 중요해졌다

2. 이론적 배경

기술 수용 이론 모델이 연구에 널리 사용됨에 따라 학자들은 모델의 설명력을 향상시키기 위해 기본 이론 모델에 새로운 변수를 지속적으로 도입하고 있다. 전통적인 전자상거래와 영향을 미치는 요인도 다르다. 이를 위해 본 논문에서는 크로스보더 전자상거래 B2C 플랫폼을 연구 대상으로 하고 기술 수용 모델을 기반으로 플랫폼 브랜드 이미지, 플랫폼 품질, 쇼핑 경험을 외부 변수로 추가하며 위험을 감지하는 중개 역할을 통해 기술 수용 모델을 확장하여 크로스보더 전자상거래 B2C 플랫폼의 소비자 구매 의향에 영향을 미치는 요인을 보다 종합적으로 분석하고 크로스보더 전자상거래 구매 의향을 예측할 수 있으며 이론적 의의가 있다.

3. 연구 설계

가. 연구 모델



(그림 1) 연구모델

나. 3.2 연구 가설

- H1. 쇼핑 경험은 지각된 유용성에 상당한 긍정적인 영향을 미칠 것이다.
- H2. 쇼핑 경험은 지각된 용이성에 상당한 긍정적인 영향을 미칠 것이다.
- H3. 쇼핑 경험은 지각된 위험에 상당한 부정적인 영향을 미칠 것이다.
- H4. 플랫폼 브랜드이미지는 지각된 유용성에 상당한 긍정적인 영향을 미칠 것이다.
- H5. 플랫폼 브랜드이미지는 지각된 용이성에 상당한 긍정

적인 영향을 미칠 것이다.

- H6. 플랫폼 브랜드이미지는 지각된 위험을 상당한 부정적인 영향을 미칠 것이다.
- H7. 플랫폼 품질은 지각된 유용성에 상당한 긍정적인 영향을미칠 것이다.
- H8.플랫폼 품질은 지각된 용이성에 상당한 긍정적인 영향을미칠 것이다.
- H9.플랫폼 품질은 지각된 위험에 상당한 부정적인 영향을 미칠 것이다.
- H10. 지각된 유용성은 구매 의도에 상당한 긍정적인 영향을 미칠 것이다.
- H11. 지각된 용이성은 구매 의도에 상당한 긍정적인 영향을 미칠 것이다.
- H12. 지각된 위험은 구매 의도에 상당한 부정적인 영향을 미칠 것이다.

4. 결 론

소비자 행동 이론과 인지적 가치 이론, 주로 크로스보더 전자 상거래 플랫폼의 관점에서 TAM 모델을 기반으로 연구 모델을 설정하고 소비자의 구매 의도에 영향을 미치는 플랫폼 요인을 분석한다. 이 글은 이전 학자들의 연구 결과와 크로스보더 전자상거래 플랫폼의 특성을 결합하여 쇼핑 경험 요인, 플랫폼 품질, 플랫폼 브랜드 이미지의 세 가지 외인변수에서 출발하여 소비자가 수익을 감지하는 지각된 유용성, 지각된 용이성, 지각된 위험이 구매 의도에 미치는 영향을 연구한다. 둘째, 설문조사를 통해 크로스보더 소비자의 표본 데이터를 수집하고 이론 모델에 대한 실증 분석을 수행한다. 쇼핑 경험, 플랫폼 품질 및 플랫폼 브랜드 이미지는 지각된 유용성에 긍정적인 영향을 미친다.쇼핑 경험, 플랫폼 품질 및 플랫폼 브랜드 이미지는 지각된 용이성에 긍정적인 영향을 미친다.쇼핑 경험, 플랫폼 품질 및 플랫폼 브랜드 이미지는 지각된 위험에 부정적인 영향을 미친다.지각된 유용성과 지각된 용이성은 구매 의도에 현저한 긍정적인 영향을 미친다. 지각된 위험을 구매 의도에 현저한 부정적인 영향을 미칠 수 있다.

참 고 문 헌

- [1] 김나민, 김해룡, 이문규, 최현국, 김일민, “인터넷 구전 수용의 영향요인.” 大韓經營學會誌, 19(4), 2006, pp.1515-1534.
- [2] 김세준, 이세진, “정보원, 메시지의 유형 및 수신자 성향에 따른 구전효과연구: 영화에 대한 온라인 비평을 중심으로.” 광고연구, 96, 2013, pp.100-136.
- [3] 김봉준, 황의록, “온라인에서의 제품 평가정보가 소비자 구매 의사결정에 미치는 영향에 관한 연구: 제품 리뷰기사와 소비자댓글의 구전효과,” 한국경영학과 통합학술대회, 2007, pp.1-27.
- [4] 김창호, 황의록, “구전정보의 특성과 구전효과의 관계, ” 광고연구, 35(여름), 1997, pp.55-77.
- [5] 김한나, “온라인 구전정보의 방향성이 구전효과에 미치는 영향에 관한 연구 : 화장품 브랜드 인지도의 차이를 중심으로,” 숙명여자대학교, 석사학위논문, 2009.

한국 전래동화 소재의 보드게임 개발을 위한 판타지적 특징 분석

이유안, 한지애¹⁾

조선대학교 디자인공학과

e-mail : liz200@naver.com, jijilion@chosun.ac.kr

Analysis of Fantastic Features for the Development of Board Games based on Korean Traditional Fairy Tales

Yu-An Lee, Ji-Ae Han

School of Informatics & Product Design, Chosun Univ.

요약

본 연구는 한국 전래동화를 소재로 한 페이퍼 프로토타입의 보드게임 개발을 위해 동화 속 판타지 요소를 분석하는 것이다. 르네 지라르(René Girard)의 ‘삼각형 욕망’ 이론과 파라 멘틀레슨의 판타지에 관한 문헌연구를 통해 욕망의 매개체, 판타지 유형, 판타지 매개체를 도출하고, 이를 보드게임 디자인 요소(목표, 서사, 매커닉스, 아트, 컴포넌트)에 적용해 페이퍼 프로토타입의 보드게임 개발을 위한 구성안을 제안했다. 본 연구의 내용 토대로 향후 다양한 매체로 구현하는 연구를 지속하고자 한다.

1. 서론

전래동화는 신화, 전설, 민담에서 발전해 이루어진 단순한 옛이야기를 넘어, 사고방식, 윤리관, 그리고 공동체 의식을 형성하는 데 중요한 역할을 해왔다. 대다수의 전래동화는 인간의 다양한 욕구와 갈등을 묘사하고 있으며, 이야기 속 문제의 실마리를 공상적이고 환상적인 장치로 풀어내면서 그로부터 도출되는 윤리적 교훈을 전달한다. 그럼에도 불구하고, 한국 전래동화를 활용한 콘텐츠 개발이 활발하지 않아 본 연구에서는 한국 전래동화를 소재로한 콘텐츠를 개발하기 위해 동화에 나타난 환상적 요소를 고찰하고 보드게임 개발을 위한 요소를 도출하고자 한다. 본 연구에서는 뱀게임, 부루마블 방식의 보드게임 방식을 착안한 페이퍼 프로토타입 기획으로 연구의 범위를 제한하고, 향후 이를 통해 콘텐츠 개발 매체를 확대하고자 한다.

2. 한국 전래동화에 나타난 욕망

한국 전래동화는 주로 권선징악, 희망, 효 등의 메시지를 담고 있으면서, 특히 강한 자/힘에 대한 약한 자의 대립이나 인간의 욕망을 주로 다루고 있다. 전래동화 속에 나타난 욕망들은 인물들의 행동과 선택을 결정짓는 중요한 요인으로 작용하며, 이로 인해 발생하는 갈등과 해결 과정을 통해 교훈을 제공한다. 표1은 문학 비평 및 철학 이론 중 하나인 르네 지라르(René Girard)의 ‘삼각형 욕망’ 이론을 기준으로 한국 전래동화에 나타난 욕망과 욕망의 주체(Subject)와 대상(Object), 욕망을 해결하는 방식인 구성요소(모방, 경쟁, 분열), 욕망의 해결을 돕는 매체를 정리한 것이다. 이론에 따르면, 욕망은 결핍과 타인의 욕

망을 모방(Mimesis)하는 데에서 오며, 그 결핍을 채우기 위해 경쟁(Rivalry)과 분열(Conflict)을 겪게 되며, 이 과정을 통해 욕망은 해소되거나 발전된다.

(표 1) 주요 한국 전래동화의 삼각형 욕망

전래동화	주체	대상	방식
홍부와 놀부	놀부	부	모방
해님달님	오누이	생존	분열
별주부전	임금	건강/치유	분열
	토끼	부/명예	경쟁/갈등
혹부리영감	욕심많은 영감	부	모방
선녀와 나무꾼	나무꾼	행복한가정	분열
도깨비 방망이	형님	부	모방
토끼와 거북이	토끼	명예	분열
심청전	심청	행복한 삶	분열
	심봉사	인간적 삶	분열
구렁덩덩 새선비	셋째딸	행복한가정	분열
	자매	행복한가정	분열

표 1에서 다룬 한국 전래동화에서의 욕망은 행복한 삶이나 평범한 가정생활, 생존, 부, 명예에 관한 것들이었으며, 욕망의 주체는 가난하고 힘 없는 평민이거나 끝없는 욕심을 ‘가진 자’였다. 혹부리 영감이나 도깨비 방망이에서 처럼 ‘가진 자’의 욕망의 중개자가 힘 없는 주인공이 되어 모방(Mimesis)으로 전개되어 권선징악 메시지를 전달하는 경우가 많다. 구렁덩덩 선비에서도 주인공인 셋째딸이 두 딸의 욕망의 중개자가 되어 분열을 초래했던 것처럼, 대다수 동화에서 욕망은 분열(Conflict)을 야기한다. 반면, 경쟁을 통한 욕망의 해소는 상대적으로 많지 않았다.

3. 한국 전래동화의 판타지적 요소

파라 멘틀레슨은 『Rhetorics of Fantasy』에서 환상이

1) 교신저자

서사에 진입하는 방식을 기준으로 판타지를 네 가지로 분류했다. 네 유형은 현실 세계와 판타지 세계를 연결하는 매개체를 통해 주인공이 새로운 세계로 이동하여 모험을 경험하는 구조인 ‘포털 판타지’, 현실과 다름없는 세계에서 판타지적 요소를 자연스럽게 경험하는 구조인 ‘몰입형 판타지’, 현실 세계에 판타지적 요소가 갑작스럽게 침입하여 사건이 전개되는 구조인 ‘틈입형 판타지’, 현실과 판타지 세계의 경계가 모호하게 그려지며, 두 세계가 자연스럽게 혼합되는 구조인 ‘리미널 판타지’이다.

표 2는 앞서 제시한 한국 전래동화의 판타지 유형, 판타지로 연결되는 서사의 매개체, 동화 속 판타지 세계(시간/공간)를 제시한 것이다. 천상, 용궁, 지하세계와 같은 판타지 세계를 현실 세계와 연결하는 포털 판타지와 리미널 판타지식 서사가 많았으며, 몰입형 판타지는 한국 전래동화에서는 찾기 힘들었다. 선녀와 나무꾼과 해님달님은 새로운 세계인 천상으로 욕망의 주체가 이동한다는 서사가 결말의 형태로 사용되어 틈입형 판타지로 보았다. 또한 욕망을 해소하는 매개체이자 판타지 세계와 연결하는 매개체는 욕망의 주체가 일상생활에서 사용하는 용품이나 현실에서 존재하지 않은 상상의 물건들이며, 이들은 욕망을 해소하거나 욕망을 해소하기 위한 비현실적 공간, 더 나아가 판타지 세계로 이끄는 매개체 역할을 담당한다.

(표 2) 한국전래동화의 판타지 유형과 특징

전래동화	판타지 유형	판타지 매개체	판타지 세계
구렁덩덩새선비	포털 판타지	뱀허물	지하세계
십정전	포털 판타지	연꽃	바다 속 용궁
별주부전	포털 판타지	-	바다 속 용궁
선녀와 나무꾼	틈입형 판타지	날개 옷	천상
해님달님	틈입형 판타지	동아줄	천상
홍부와 놀부	틈입형 판타지	박씨	-
도깨비 방망이	리미널 판타지	도깨비방망이	-
혹부리영감	리미널 판타지	도깨비방망이	-

4. 판타지적 요소를 활용한 보드게임 디자인 요소

보드게임의 디자인 요소인 스토리, 메커닉스, 컴포넌트, 아트, 목표와 각 요소의 하위 요소(이은정, 2024)를 기준으로 표3과 같이 한국 전래동화의 판타지 요소를 적용했다. 표 2에서 제시한 전래동화의 판타지적 특징을 반영해 보드게임의 공간적 요소를 천상, 바다 속 세계, 지하세계, 판타지 요소가 있는 현실 세계로 설정하고, 네 개의 세계를 플레이어가 여행하는 로드맵 방식의 ‘Multi-Universe’ 형식의 세계관으로 구성했다. 각 세계와 플레이어를 연결하는 판타지 매개체는 각 동화의 서사에 중요한 사물이면서 판타지 세계로 연결해 주는 매개체를 환유적 시각 모티브로 설정했다. 네 개의 판타지 공간의 로드맵에 동아줄, 도깨비 방망이와 같은 판타지 매개체를 배치해 전래동화의 메시지나 내용을 암시하면서 보드게임의 규칙을 표시하는 컴포넌트 역할을 담당하도록 했다.

본 콘텐츠는 욕망의 세계들을 지나 목적지에 도착하는

목표와 단위 퀘스트로 구성된다. 단위 퀘스트는 각 동화의 판타지 매개체를 만나 스토리를 보고, 보상을 받고(판타지 매개체) 그에 따라 지시하는 곳으로 이동하는 방식이다. 그래픽 컨셉은 옴아트 형식의 로드맵과 리미널 판타지 형식을 구성하기 위한 네 개의 세계를 비현실적으로 합성하는 방식으로 구현했다.

(표 3) 한국전래동화의 판타지 요소 게임요소

목표	미션	각 NPC들의 욕망의 세계를 지나 목적지 도착
	퀘스트	각 동화 속 욕망의 주체를 만나 아이템 얻기
서사	세계관	네개의 세계/Multi-verse, 동화 속 욕망이 하나로 어우러진 공간
	스토리	네 개의 세계 속 동화를 만나는 여행
	캐릭터	동화의 욕망의 주체들이 NPC
메커닉스	규칙	뱀게임과 부르마블의 규칙 사용
	보상	각 동화의 스토리에 부합해 +, -
아트	옴아트 형식의 로드맵, 리미널 판타지형식의 비현실적 그래픽	
컴포넌트	보드, 스토리카드, 주사위, 게임말, 보상카드	



(그림 1) 판타지적 요소를 활용한 한국 전래동화 보드게임 컨셉안

5. 결론

본 연구는 한국 전래동화의 문화적 가치를 현대적인 보드게임 형식으로 재해석함으로써 전통문화의 새로운 진흥 방식을 제시한다는 점에서 의미가 있다. 한국 전래동화는 단순한 오락적 이야기 이상의 의미를 지니며, 한국인의 정체성과 가치관을 형성하는 데 중요한 역할을 해왔다. 그러나 현대 사회의 급속한 변화 속에서 이러한 전통적 이야기가 잊혀지거나 그 중요성이 약화되고 있는 상황에서, 본 연구는 전래동화를 현대적 미디어 형식인 보드게임으로 재구성하여 그 가치를 새롭게 조명했다. 본 연구에서는 페이퍼 프로토타입 방식으로 한국 전래동화를 현대적으로 재구성하기 위한 기초연구에 집중했으나, 본 연구를 토대로 다양한 매체로 구현하는 연구를 지속하고자 한다.

참고 문헌

- [1] 서기환, “현대인의 욕망에 대한 판타지적 표현연구 - 본인의 작품과 교육사례를 중심으로”, 『미술문화연구』 제28권 제28호, 2024. pp.63-95.
- [2] 이은정, 『게임디자인&페이퍼 프로토타입』, 길벗캠퍼스, 2024. pp. 144-152

게이미피케이션 교육콘텐츠에서 교육내용-전략-게임요소 관계

한지애
 조선대학교 디자인공학과
 e-mail : jijilion@chosun.ac.kr

Relationship of Contents-Strategy-Game Elements in Gamification Educational Contents

Ji-Ae Han
 School of Informatics & Product Design, Chosun Univ.

요약

본 연구는 게이미피케이션 교육콘텐츠에서 교육내용과 게임요소의 관계에 관한 고찰이다. TPACK 모형의 내용지식, 교육지식, 기술지식의 개념을 토대로 게이미피케이션 교육 콘텐츠 설계 개념을 고찰했다. 내용지식은 메릴(Merill)의 내용요소 제시이론에서, 교육지식은 스키펴딩 전략에서, 기술지식은 게이미피케이션 교육 콘텐츠 개발 관련 선행 연구를 토대로 개념화했으며, 이를 토대로 게이미피케이션 교육 콘텐츠 개발을 위한 내용 수준별 주요 교육전략과 게임요소를 제안하였다.

1. 서론

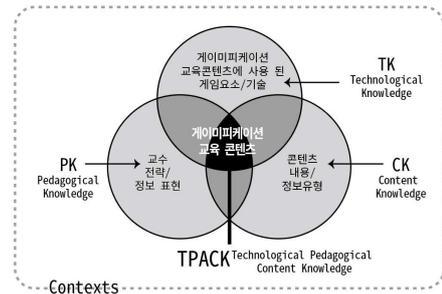
Simões(2013)은 게임 요소를 적용한 수업 설계를 위한 가이드라인으로 목표 달성을 위한 반복, 즉각적인 피드백, 수준에 맞는 과제 제공과 난이도 조절, 다양성과 보상을 제시했다. 모든 가이드라인이 학습 목표와 목표에 따른 교육 방법에 초점을 맞추고 있다. 그러나 게임 요소를 적용한 대다수의 교육 콘텐츠 개발에 관한 연구는 교수법 또는 게임 요소를 적용한 개발 방법론에 집중하고 있다. 이에 본 연구는 TPACK(Technological Pedagogical and Content Knowledge) 모형을 토대로 교육 내용, 교수법과 게임 요소 적용이라는 세 가지 관점에서 게이미피케이션 교육 콘텐츠에서 교육 내용과 게임 요소의 관계를 살펴보고자 한다.

2. TPACK 모형기반의 게이미피케이션 교육콘텐츠의 설계 개념

Shulman(1987)에 미국교육의 문제점을 지적하며 소개한 TPACK 모형은 효과적으로 기술을 교육에 도입해 교육 목표와 내용에 부합한 교육방법을 제시하는데 목적을 두고 있다. 모형에 따르면 교수자는 교수설계 시 세 가지 핵심요소인 기술적 지식(Technological Knowledge), 교육적 지식(Pedagogical Knowledge), 내용 지식(Content Knowledge)을 적절하게 결합해 수업을 설계해야 한다.

TPACK 모형을 토대로 게이미피케이션 교육콘텐츠 개발의 교육학적 개념을 그림1과 같이 제시하였다. 게이미피케이션 교육콘텐츠 디자이너는 교과내용에 관한 내용지식(CK)을 교육학적 지식(PK)를 토대로 정보디자인 방법으로 구성하고, 게임 디자인 요소(게임에 관한 지식_TK)를 교육목표에 맞게 구성해야 한다. 따라서, 게이미피케이션

교육 콘텐츠의 설계는 교육 내용에 부합한 교육 전략으로써 게임 요소를 활용하는 것에 초점을 맞추고 있다.



(그림 1) TPACK 기반의 게이미피케이션 교육콘텐츠 개념

3. 내용지식을 위한 내용요소 제시이론(CDT)

교수설계이론은 교육 목표 달성을 위한 방법을 계획하는데 초점을 두고 있다. 교수(教授) 방법은 ‘조직전략’, ‘전달 전략’, ‘관리전략’을 포함하며, 본 연구에서는 게이미피케이션 교육 콘텐츠 개발을 위한 것으로 ‘조직전략’에 초점을 두었다. 거시적 조직전략에 부르너(Bruner)의 나선형 교육 과정과 레이겔루스(Reigeluth)의 정교화이론이 있으며, 미시적 조직전략에는 Merrill의 내용 요소제시 이론(Component Display Theory: CDT)이 대표적이다. 그는 네 가지 교육 내용과 세 가지 수행 수준으로 교육내용에 관한 분류체계를 만들고, 학습 결과를 유도하기 위해 두 단계의 제시형태(교육 방법)를 제안했다. 내용 요소 제시이론의 핵심인 ‘수행×내용 매트릭스’는 네 가지 내용 형태(사실, 개념, 절차, 원리)를 기억, 활용, 발견하는 수행 수준으로 나누어 교육하는 방식이다. 기억단계에 해당하는 교육 내용들은 주로 암기를 통한 교육 내용을 습득하는 수준으로 학습 내용이 감각 저장고에 저장되는 부분에 해

당한다. 활용단계는 단기 저장고에서 수행을 통해서 교육 내용이 인지되고, 발견단계는 학습자가 학습을 통해 새로운 내용을 창안하는 인지 전략 수준에 해당한다.

표 1은 내용요소 제시이론의 네 가지 내용수준(사실, 개념, 절차, 원리)의 정의와 정보디자인 측면에서 정보특징과 지식의 유형(Peter Durcher가 제안한 다섯 가지 지식 유형)을 제시한 것이다. 사실과 개념 내용은 명시적 지식으로 학습 대상(What) 자체에 관한 것이며, 절차적 내용은 순차적 방법에 관한 지식으로 구체적인 상황에 맞는 방법(How)에 관한 것이다. 원리적 내용은 인과관계에 관한 것으로 원인(Why)에 대한 규명과 결과 활용에 관한 것이다.

(표 1) 내용요소 제시이론의 네 가지 내용수준

내용수준	정의	정보특징	지식의 유형
사실 (Fact)	사물, 사건, 장소 이름과 같은 선언적 내용	선언적	명시적 지식 Know-What
개념 (Concept)	공통적 속성을 지닌 사물, 사건들의 집합	설명적	
절차 (Procedure)	특정 목표나 해결을 위한 일련의 단계들 또는 순서	순차적	방법적 지식 Know-How
원리 (Principle)	개념들 사이의 관계나 상호관련성 정의	인과관계	인과적 지식 Know-Why

4. 교육적 지식을 위한 CDT와 스캐폴딩 전략

내용 요소제시 이론에서 교육 목표와 교수 방법에 맞게 교과 내용을 제시하는 형태를 통칭하는 개념을 ‘자료제시형’이라 한다. 일차자료제시형은 최소한의 제시방식으로, 말로 알려주는 형식의 설명식(Expository), 질문과 수행형식의 탐구식(Inquisitory), 일반성(Generality), 구체적 예시인 사례(Instance)로 구분된다. 이외에 부가적인 자료제시 형태인 ‘이차자료제시형’으로 학습의 효과를 높여주기도 하며, 일차자료제시형과 함께 사용되어 내용 전달을 정교화시킨다.

(표 2) 내용 전달 방식과 관련 스캐폴딩 주요 전략

이차 자료제시형	정의	관련 스캐폴딩 전략
맥락(Context)	내용과 관계 있는 맥락이나 역사적인 배경 제시	구체화
선수학습 (Prerequisites)	새로운 학습 내용 관련 선수지식 제시	배경지식활용
기억술 (Mnemonic)	익숙한 정보로 새로운 정보를 기억하게 하는 방법제시	중요특성 표시하기
의미정교화의 도움 (Focusing Helps)	내용에 집중할 수 있는 시각 요소 사용	안내, 조인
표상법 (Representation)	학습내용을 개념화 된 그림 또는 인포그래픽으로 제시	과제 구조화
피드백(Feedback)	평가와 피드백 제공	피드백, 평가

스캐폴딩 전략은 학습자가 스스로 학습능력을 향상할 수 있도록 교수자가 지원해 주는 교육방법으로, 인지적 스캐폴딩과 정서적 스캐폴딩 또는 거시적 스캐폴딩과 미시적 스캐폴딩으로 나뉜다. 본 연구에서는 인지적·미시적 스캐폴딩 전략을 중심으로 표2와 같이 이차 자료제시형의

정의와 이와 관련된 스캐폴딩 전략을 제시하였다.

5. 내용지식에 따른 교육전략과 게임 요소

기술지식(TK)은 테크놀로지를 효과적으로 사용하기 위한 이론적·기술적 지식으로, 본 연구에서는 교육컨텐츠 개발을 위한 게임적 지식에 해당한다. 교육학적-내용지식을 위한 기술지식으로서 게이미피케이션 교육 콘텐츠의 세 가지 게임 요소는 목표(미션/퀘스트, 서사), 보상(진행적·사회적 보상), 프로세스(피드백, Mechanic Loop, 플레이 구도, 관계/참여)이다.(한지애, 2022) 앞서 제시한 네 가지 내용지식 유형과 그에 따른 주요 교육 전략에 따른 주요 게임 요소를 표 3과 같이 제시하였다. 사실적 내용과 개념적 내용은 상대적으로 일차자료제시형 중심의 교육전략이 주로 사용되며, 콘텐츠에서는 스토리, 세계관을 배경으로 캐릭터 대사나 Help장이 주로 사용된다. 절차적 내용과 원리적 내용은 상대적으로 이차자료제시형 중심의 교육전략이 주로 사용되며, 진행적 보상과 Mechanic Loop가 주로 콘텐츠에서 사용된다.

(표 3) 내용 수준별 주요 교육전략과 게임요소

내용	주요 교육전략	주요 게임 요소
사실	설명, 조직&연계, 반복, 기억술, 맥락	스토리, 맵(공간), 정보창, 캐릭터 대사/대화
개념	비유, 사례, 속성 분류, 선수학습	스토리, 세계관, 단위 퀘스트, Help, 커스터마이징
절차	개요, 설명, 시연, 의미 정교화의 도움, 피드백	Play Guide, Mechanic Loop, 아이템, 부스터, Help
원리	표상, 설명, 선수학습, 피드백, 맥락	설명, Mechanic Loop, 아이템, 부스터, Help

6. 결론

본 연구는 게임 요소를 적용한 교육컨텐츠 설계 가이드라인을 개발하기 위한 기초연구로, 교육학적인 측면에서 교육내용과 콘텐츠화 방법(교육전략과 게임요소)의 관계에 관한 고찰로 진행되었다. 본 연구에서는 TPACK 모형에 기반 해 내용 수준별 주요 교육 전략과 게임 요소를 제안하는데 그쳤으나, 본 연구를 토대로 교육목표-정보특성-퀘스트-게임요소를 제안하는 개발 프레임워크를 제안하고자 한다.

“이 논문은 2023년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구지원사업의 지원을 받아 수행된 연구임(NRF-2023S1A5A2A01080111)”

참고문헌

- [1] 이선희, “고등교육을 위한 게임화 수업설계모형 및 지침개발” 충북대학교, 박사학위논문, 2021
- [2] 백영근 외, 『스마트 시대의 교육방법 및 교육공학』, 학지사, 2015.
- [3] 한지애, “게이미피케이션 교육 콘텐츠 디자인을 위한 정보 조직화 모델” 한국과학예술융합학회, 제40권 제5호, 2022. pp.439-452.

특허정보를 활용한 국내 반려동물 산업 동향에 관한 연구: 네트워크 분석과 텍스트 마이닝 기법을 활용하여

이동현, 강주영, 이한솔
아주대학교 비즈니스애널리틱스학과
swagger910@ajou.ac.kr, jykang@ajou.ac.kr, lhs15@ajou.ac.kr

A study on trends in the domestic pet industry using patent information: Using network analysis and text mining

Donghyeon Lee, Juyoung Kang, Hansol Lee
Department of Business Analytics, Ajou University

요 약

반려동물의 가족화 현상으로 인하여 반려동물 관련 산업이 크게 성장하고 있으며 미래 핵심 산업으로 주목받고 있으나, 국내 반려동물 산업에 대한 특허 빅데이터를 활용한 연구는 아직 부족하다. 이에 본 연구에서는 반려동물 산업의 동향과 기술 융합 양상을 파악하기 위한 세 가지 분석을 진행하였다. 첫 번째로, 웹스온(wipson)에서 검색식을 활용하여 특허 데이터를 수집한 후, IPC (International Patent Classification) 특허 분류 코드를 활용하여 반려동물 관련 기술의 융합 발전 양상을 분석하였다. 두 번째로 반려동물의 입양이 크게 증가한 시점인 2020년을 기준으로 2015~2019년, 2020~2024년으로 5개년씩 기간을 나눈 후 분류 중심성 척도를 기준으로 출원된 반려동물 관련 특허 분류 코드를 대조하여 융합 발전 양상을 비교하였다. 세 번째로 각 기간의 반려동물 특허의 내용을 분석하여 반려동물 기술과 관련된 토픽들을 도출하였다. 분석 결과, 2020년 이전 5개년도의 경우 B23B로 시작하는 특허의 비율이 70%였으나 2020년 이후에는 40%로 감소하였으며, 2020년 이전에는 소재 및 반려동물 용품 관련 특허 출원이 주를 이룬 반면, 2020년 이후에는 반려동물 기능성 식품의 제조에 관련된 특허가 상위에 등장하였다. 이후 특허 데이터의 ‘발명의 명칭’을 대상으로 토픽 모델링을 진행하였다. LDA 토픽 모델링 분석과 Perplexity를 활용하였으며, ChatGPT4o를 활용하여 토픽의 명칭에 대한 후보를 추출하고 이를 선정하였다. 토픽이 도출된 후, ChatGPT4o를 활용하여 토픽을 구성하는 키워드의 유사도가 높은 토픽들을 통합하여 최종 토픽 리스트를 도출하였다.

1. 서 론

반려동물 산업의 동향과 융합 양상을 분석하기 위하여 본 연구에서는 세 가지 연구주제를 도출하였다.

연구주제 1) 어떤 기술 분야가 반려동물 관련 기술과 결합 및 융합되어 발전하고 있는가?

연구주제 2) 국내 반려동물 특허의 세부 기술 구성은 무엇으로 구성되어 있는가?

연구주제 3) 팬데믹 이전과 이후 반려동물 특허 네트워크를 비교했을 때, 어떤 차이가 있는가?

본 연구에서는 연구주제에 대한 답변을 제시하기 위해 반려동물 관련 기술의 융합 발전 패턴을 분석하고자 한다.

2. 관련 연구

본 연구에서는 기술 발전 패턴의 파악을 위해 두 가지 분석을 진행한다. 첫 번째 분석은 IPC (International Patent Classification) 특허 분류 코드를 활용한 융합 발전 양상 분석이다. 모든 특허에는 1개 이상의 IPC 코드가 부여되며, 다양한 기술이 결합된 경우 2개 이상의 IPC 코드가 부여된다. 두 번째 분석은 2020년 이전 5개년도와 2020년 이후 5개년도의 융합 발전 양상의 대조이다. 2020년부터

시작된 팬데믹 이후 반려동물의 입양이 크게 증가하며, 특허 출원 양상에도 변화가 발생했다. 본 연구에서는 기술 네트워크 분석을 통해 반려동물 관련 산업의 발전 패턴의 변화 양상을 시각화하고 분석한다. 세 번째 분석은 반려동물 특허의 내용을 분석하여 반려동물 기술과 관련된 토픽들을 도출한다.

참 고 문 헌

- [1] 김수정, “반려동물 시장 8조원 규모… 바야흐로 펫코노미(Pet + Economy) 시대,” 조선일보, 2023.08.30.
- [2] 송중호, ““2027년까지 2배 확대 15조로”…펫시장, 전략산업으로 키운다,” 서울경제, 2023.08.09, <https://www.sedaily.com/NewsView/29TCAXV90>.
- [3] 최지희, 박은정, 이혜정, “국내 반려동물 식품 및 용품 시장현황 분석 연구,” 한국콘텐츠학회논문지, 제19권, 제8호, 115-122쪽, 2019년 8월.
- [4] 이정민, 장민아, “반려동물 용품 및 화장품 산업 현황,” 문화기술의 융합, 제7권, 제4호, 833-844쪽, 2021년 11월.
- [5] 이윤주, 송준석, “펫푸드(반려동물 식품)분야 국내 특

허 동향 분석,” 문화기술의 융합, 제9권, 제5호, 625-633쪽, 2023년 9월.

[6] 김세형 외, “온라인 PC 게임과 모바일 게임의 기술 발전 패턴 분석 및 전략 수립 연구: 네트워크 분석과 텍스트 마이닝 기법을 활용하여,” 2023.

[7] 한수정, “반려동물 동반관광의 관광동기, 관광제약, 행동의도 간의 관계 연구,” 호텔리조트연구, 제21권, 제5호, 217-234쪽, 2023년 12월.

[8] 김미경, 김지은, “반려동물 동반관광에 관한 탐색적 연구,” Tourism Research, 제45권, 제4호, 19-38쪽, 2020년 12월.

[9] 강정구, “팻캠족의 출현과 반려동물의 재인식 - 2000년대 이후 한국문학과 영화 작품을 중심으로,” 세계문학비교연구, 제54권, 5-27쪽, 2016년.

[10] Clive-Steven Curran, Stefanie Bröring, Jens Leker, “Anticipating converging industries using publicly available data,” Technological Forecasting and Social Change, vol. 77, no. 3, pp. 385-395, 2010.

[11] Y. Cho, E. Kim, and W. Kim, “Strategy transformation under technological convergence: Evidence from the printed electronics industry,” International Journal of Technology Management, vol. 674, pp. 106-131, 2015.

[12] Clive Curran and Jens Leker, “Patent indicators for monitoring convergence - examples from NFF and ICT,” Technological Forecasting and Social Change, vol. 78, no. 2, pp. 256-273, 2011.

[13] M. Garcia-Murillo and I. MacInnes, “The impact of technological convergence on the regulation of ICT industries,” Journal of Media Management, vol. 5, no. 1, pp. 57-67, 2003.

[14] O. Kwon, Y. An, M. Kim, and C. Lee, “Anticipating technology-driven industry convergence: evidence from large-scale patent analysis,” Technology Analysis & Strategic Management, vol. 32, no. 1, pp. 1-16, 2019.

[15] S. Chae and J. Gim, “A Study on Trend Analysis of Applicants Based on Patent Classification Systems,” Information, vol. 10, no. 12, p. 364, 2019.

[16] K. Lee and B. Yoon, “Developing an research and development (R&D) process improvement system to simulate the performance of R&D activities,” Computers in Industry, vol. 92-93, pp. 178-193, 2017.

[17] C. Lee and G. Lee, “Technology opportunity analysis based on recombinant search: patent landscape analysis for idea generation,” Scientometrics, vol. 121, no. 2, pp. 603-632, 2019.

[18] H. Park, T. Anderson, and W. Seo, “Regional

innovation capability from a technology-oriented perspective: An analysis at industry level,” Computers in Industry, vol. 129, p. 103441, 2021.

[19] O. Kwon, Y. An, M. Kim, and C. Lee, “Anticipating technology-driven industry convergence: evidence from large-scale patent analysis,” Technology Analysis & Strategic Management, vol. 32, no. 1, pp. 1-16, 2019.

[20] 박수현, 윤영미, 김호용, 김재수, “5개국 바이오헬스 산업의 기술융합과 트렌드 분석: 특히 동시분류분석과 텍스트마이닝을 활용하여,” 한국융합학회논문지, 제12권, 제4호, 9-21쪽, 2021년 4월.

[21] 고재창, 조근태, 조윤호, “키워드 네트워크 분석을 통해 살펴본 기술경영의 최근 연구동향,” 지능정보연구, 제19권, 제2호, 101-123쪽, 2013년 6월.

[22] 조재인, “네트워크 텍스트 분석을 통한 문헌정보학 최근 연구 경향 분석,” 정보관리학회지, 제28권, 제4호, 65-83쪽, 2011년 12월.

[23] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” Journal of machine Learning research, vol. 3, no. 1, pp. 993-1022, 2003.

[24] U. Chauhan and A. Shah, “Topic modeling using latent Dirichlet allocation: A survey,” ACM Computing Surveys (CSUR), vol. 54, no. 7, pp. 1-35, 2021.

[25] 이봉우, “상세 서지 분석과 토픽 모델링을 이용한 발명 교육 연구의 국제 동향 분석,” 교과교육학연구, 제27권, 제1호, 1-12쪽, 2023년 2월.

[26] M. E. Emenike and B. U. Emenike, “Was this title generated by ChatGPT? Consideration for artificial intelligence text-generation software programs for chemists and chemistry educators,” Journal of Chemical Education, vol. 100, no. 4, pp. 1413-1418, 2023.

[27] C. Chatfield, “Exploratory data analysis,” European journal of operational research, vol. 23, no. 1, pp. 5-13, 1986.

글로벌 전자상거래가 중국의 수출에 미치는 영향

제1저자 : XIA ZUOHAO(하작호)
 전남대학교 대학원 디지털미래융합서비스협동과정
 e-mail : 1020275136@qq.com
 교신저자 : 박복재
 전남대학교 글로벌비즈니스학부

The Impact of Global E-Commerce on China's Exports

DIGITAL FUTURE CONVERGENCE SERVICE
 CHONNAM NATIONAL UNIVERSITY

요 약

코로나19의 세계적 대유행, 국제 경제력 균형의 중대한 변화, 경제 세계화의 쇠퇴, 국제시장 수요 둔화 등으로 인해 중국의 전통 무역 성장은 영향을 받았다. 국제무역의 새로운 모델인 글로벌 전자상거래는 디지털 무역의 필수적인 부분이기도 하며 급속한 발전 과정에서 지속적으로 개선되고 표준화되었으며 점차 중국 대외 무역 발전의 강력한 출발점이 되었다. 따라서 글로벌 전자상거래 발전과 중국 수출의 상호작용 관계를 연구하는 것은 중요한 실무적 의의를 갖는다.
 키워드: 글로벌 전자상거래, 수출, 중국

1. 서 론

1940년대 이후 컴퓨터 기술로 대표되는 3차 산업혁명이 등장하면서 경제, 사회 활동 전반에 큰 영향을 미치고 있다. 국제무역의 발전을 살펴보면, 전통적인 최종 제품 무역과 글로벌 가치 사슬 무역 단계를 거친 후 기술 변화는 국제무역의 디지털 무역으로의 전환을 촉진하고 있다. 단계에서는 무역 활동이 현대 정보 네트워크를 사용하여 제품과 서비스를 교환함으로써 소비자 인터넷을 산업 인터넷으로 업그레이드하고 궁극적으로 새로운 유형의 지능형 제조 무역으로 업그레이드하도록 만든다.

글로벌 전자상거래의 발전이 증가함에 따라 학계에서도 글로벌 전자상거래에 관한 학술 연구가 늘어나고 있다. 글로벌 전자상거래에 대한 현재 학술 연구는 주로 영향 요인, 무역 모델, 무역 성장과의 상호 관계라는 세 가지 측면에 중점을 두고 있다.

인터넷과 대외 무역의 결합으로 글로벌 전자상거래의 새로운 무역 모델이 탄생했다. 경제 성장의 "트로이카" 중 하나로서 수출은 중국과 같은 신흥 경제국에서 중요한 역할을 한다. 글로벌 전자상거래가 중국의 수출에 어떤 영향을 미칠 것이며 구체적인 영향 경로 메커니즘은 무엇입니까? 이는 중국과 글로벌 전자상거래에 참여하는 모든 경제 주체가 고려해야 할 문제이다. 본 논문은 이를 바탕으로 구체적인 분석을 진행하는데, 이는 풍부한 이론적, 실제적 의의를 지닌다.

글로벌 전자상거래는 전통적인 대외무역에 새로운 생산력을 더하며, 이는 정보화와 현대화 발전의 '가속기' 역할을 할 수 있다. 특히 전자 도구의 발전은 인간 노동의 역사를 변화시켰으며, 이제는 가상 공간에서의 상품과 서비스도 중요한 생산 대상이 되었다. 모바일 인터넷과 스마트폰의 확산은 글로벌 전자상거래의 성장을 촉진하고 있다.

인터넷과 스마트 기기의 보급으로 생활과 쇼핑 방식이

변화하고 있다. 글로벌 전자상거래는 중국의 국제무역 지위를 향상시키고, 무역 마찰과 제정 문제를 해결하는 데 기여할 수 있다. 비용 측면에서 유리한 전자상거래는 향후 국제무역의 중요한 동력이 될 것이다. 하지만 물류, 통관, 결제 등 해결해야 할 문제들도 남아 있다. 빅데이터와 클라우드 컴퓨팅의 발전은 전자상거래 산업의 성장을 가속화하고 있으며, 이는 무역 효율성을 높이는 기회로 활용되어야 한다.

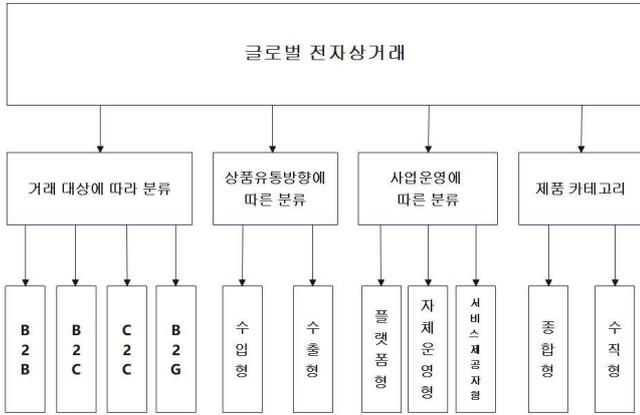
2. 관련 연구

본 논문은 Alibaba 글로벌 전자상거래 연구센터의 글로벌 전자상거래 정의를 따릅니다. 즉, 글로벌 전자상거래는 국제 전자상거래 마케팅 네트워크를 사용하고 거래를 완료하고 결제를 수행하는 전자상거래 플랫폼이다. 국경 간 물류를 통해 결제, 상품 배송 및 거래 완료를 포함하는 국제 비즈니스 활동이다.

전자상거래는 처음에는 B2B와 B2C의 두 가지 유형이 있었고 이후에는 기업이 운영하는 상품 범주에 따라 B2B, B2C, B2G, C2G의 네 가지 유형이 파생되었다. 기업이 운영하는 상품의 종류에 따라 글로벌 전자상거래는 수직형과 포괄적형으로 나눌 수 있으며, 사업 운영 유형에 따라 플랫폼형과 자체 운영형으로 나눌 수 있다. 馬述忠 등(2018)은 글로벌 전자상거래를 제3자 플랫폼, 제3자 플랫폼을 통해 거래하는 기업, 서비스 제공자라는 세 가지 범주로 분류한다. 동시에 글로벌 전자상거래 상품 무역의 대부분은 서로 다른 국가의 지리적 범위를 거쳐 이루어지기 때문에 글로벌 전자상거래는 상품 유통 방향에 따라 수출형과 수입형으로 나눌 수 있다.

본 논문에서는 글로벌 전자상거래를 관련 연구를 바탕으로 4가지

방향으로 간단하게 분류하였으며, 거래 대상에 따라 상품 유통 방향에 따라 B2B, B2C, C2C, B2G의 네 가지 범주로 구분하였다. 수입형과 수출형으로 나누어지며, 사업운영에 따라 플랫폼형, 자체운영형, 서비스 제공자형 등 4가지로 구분되며, 제품 카테고리에 따라 종합형과 수직형으로 구분된다.



(그림 1) 글로벌 전자상거래 분류

3. 연구설계과 연구가설

3.1 연구설계

Williamson(1985)의 거래비용 이론에 따르면, 거래비용은 인간의 한계와 거래 환경의 상호 작용으로 인해 발생하며, 이는 시장 실패의 원인이 될 수 있다. 글로벌 전자상거래에서는 네트워크 플랫폼을 통해 정보를 교환하여 정보 검색 비용을 절감하고, 결과적으로 거래 비용을 줄일 수 있습니다. 이는 기업과 소비자가 글로벌 전자상거래 플랫폼을 통해 연결되며, 불확실성을 감소시키고 시장의 효율성을 높이는 방식으로 작동한다.

무역 거리 관점에서, 전통적인 국제무역에서는 지리적 거리가 커질수록 운송 비용과 시간이 증가하여 수출에 부정적인 영향을 미친다. 반면, 글로벌 전자상거래는 우편, 국제 택배, 해외 창고를 통해 제품을 운송하며, 지리적 거리의 영향은 있지만, 인터넷을 통해 직접적인 소통과 거래가 이루어지므로 그 부정적인 영향을 어느 정도 완화할 수 있다. 특히 정보 검색 비용, 의사소통 장애 비용, 계약서 작성 및 이행 비용 등 정보 왜곡과 관련된 거래 비용이 크게 감소한다.

경제 규모 관점에서는, 규모의 경제가 생산 규모가 커질수록 평균 비용이 감소하는 현상으로, 이는 글로벌 전자상거래 플랫폼에서도 동일하게 적용된다. Metcalfe의 법칙에 따르면 네트워크의 가치는 네트워크에 참여하는 노드 수의 제곱에 비례하여 증가한다. 즉, 글로벌 전자상거래 플랫폼의 사용자 수가 늘어날수록 공급자에게 이익이 커지며, 이러한 플랫폼은 네트워크 외부성을 갖는다. 이는 네트워크의 규모가 커질수록 공급자와 수요자 모두에게 긍정적인 효과를 가져다준다는 것을 의미한다.

따라서 글로벌 전자상거래는 거래비용의 절감, 지리적 거리의 부정적 영향 감소, 규모의 경제 효과 등을 통해 전통적인 국제무역과는 다른 방식으로 무역 환경을 혁신하고 있다.

본 논문은 글로벌 전자상거래 수준 측정 시스템을 구축하기 위해 세계 은행 연구원인 Wilson et al.의 방법을 활용한다. 대표성, 포괄성, 비교가능성의 원칙을 바탕으로 지표를 선정하고, 다지수 차원축소법을 사용해 측정한다.

본 논문에서는 정보기술 서비스, 인프라 구축, 제도환경, 관세관리

등 4개의 1차 지표를 선정하고, 각 1차 지표의 하위 항목을 하나씩 개선하여 세계은행에서 16개의 2차 지표를 선정한다. 최종적으로 표과 같은 글로벌 전자상거래 발전 수준 측정 시스템이 구축되었다. 2013~2022년 중국의 누적 무역 수출액 상위 50개국을 선정해 T=10, N=50 패널 데이터 세트를 구축·구축되었다. 무역 상대국 간의 무역 데이터를 설명 변수로 하여, 글로벌 전자상거래 수준을 대표하는 지표를 설명된 변수로 한다.

문헌 연구와 실증 분석의 효율성을 더욱 검증하기 위해 글로벌 전자상거래와 수출무역 간 관계를 탐색한 사례로 엔커(Enker Innovations) 회사를 선정해 분석했다.

3.2 연구가설

(표 1) (연구가설)

가설번호	가설 내용
H1	글로벌 전자상거래 수준이 높을수록 중국의 수출액이 커집니다.
H2	글로벌 전자상거래 수준은 거래 비용을 통해 수출 무역에 영향을 미칩니다. 글로벌 전자상거래의 거래 수준이 높을수록 거래 비용이 낮아지고 우리나라의 수출액이 커집니다.

4. 결론

중국 수출 무역의 경우, 글로벌 전자상거래 규모는 긍정적인 영향을 미칠 수 있다.

글로벌 전자상거래는 무역 촉진 수준을 높이고 저렴한 온라인 무역 대리인을 구축하며 무역 연결 고리를 줄이고 무역 위험이 가져오는 부정적인 영향을 극복하여 중국의 무역 수준을 향상시킬 수 있다.

거래 비용은 글로벌 전자상거래의 중국 수출 무역 촉진에 상당한 영향을 미칠 수 있다. 글로벌 전자상거래는 네트워크 플랫폼을 기반으로 정보를 교환하기 때문에 거래 비용을 어느 정도 절감할 수 있다. 즉, 한 국가의 크로스보더 전자상거래 수준이 높아지면 거래비용이 절감되어 수출 교역량이 증가하게 된다.

참고 문헌

- [1] Ma Shuzhong, Fang Chao, Liang Yinfeng. Digital trade and its contemporary value and research prospects [J]. International Trade Issues, 2018(10): 16-30.
- [2] Ma Shuzhong, Fang Chao. Cross-border e-commerce and China's new export growth - based on the dual perspective of information costs and economies of scale [J]. Economic Research, 2021, 56(06): 159-176.
- [3] Zhu Xianqiang. Research on the impact of cross-border e-commerce on China's import and export trade[D].PhD thesis.University of Foreign Economics and Trade, 2020.

KOC(키 오피니언 소비자)특성이 구매의도에 미치는 영향 및 소비자 성별의 조절효과 -샤오홍수를 중심으로-

제 1 저자: CHU WEICHENG (저위성)

전남대학교 대학원
디지털미래융합서비스협동과정
e-mail : 2285664545@qq.com
교신저자: 박복재
전남대학교 글로벌비즈니스학부
e-mail: bjpark73@jnu.ac.kr

The Impact of Key Opinion Consumer (KOC) Characteristics on Purchase Intentions and the Moderating Effect of Consumer Gender -Focused on Xiaohongshu-

CHU WEICHENG Bok-Jae, Park
Chonnam National University
Digital Future Convergence Service Cooperative Course

요 약

인터넷과 소셜 이커머스의 발전에 따라 KOC(키 오피니언 소비자, Key Opinion Consumer)는 소비자의 구매 행동에 중요한 영향을 미치는 요인으로 자리 잡았다. 본 연구는 샤오홍수 플랫폼을 배경으로, S-O-R 모델(Stimulus-Organism-Response)과 Mayer의 신뢰 모델을 바탕으로 KOC 특성(전문성, 제품 관여도, 동질성)이 소비자의 지각된 신뢰성을 매개로 구매 의도에 미치는 영향을 분석하였으며, 성별의 조절 효과 또한 탐구하였다. 이를 위해 샤오홍수 사용자 363명의 유효 데이터를 대상으로 SPSS 26.0과 AMOS를 활용하여 다양한 분석을 실시하고 가설을 검증하였다.

연구 결과, KOC의 전문성, 제품 관여도, 동질성은 모두 소비자의 지각된 신뢰성과 구매 의도에 긍정적인 영향을 미쳤다. 특히 동질성이 가장 큰 영향을 미쳤으며, 지각된 신뢰성은 KOC 특성과 구매 의도 간의 관계에서 중요한 매개 역할을 하였다. 성별에 따른 분석 결과, 남성 소비자는 KOC의 동질성에 더 민감하게 반응한 반면, 여성 소비자는 KOC의 전문성과 신뢰성을 더 중요하게 고려하는 경향을 보였다. 또한 남성은 소품 동영상 리뷰와 라이브 커머스에 더 큰 반응을 보였고, 여성은 롱폼 동영상 리뷰와 이미지 및 텍스트 리뷰에 대한 신뢰도가 높았다.

이러한 결과를 바탕으로, 본 연구는 브랜드가 성별에 따라 차별화된 KOC 마케팅 전략을 수립할 필요가 있음을 제안하며, 이를 통해 소비자 신뢰를 구축하고 구매 의도를 효과적으로 높일 수 있을 것이다.

키워드: KOC,인플루언서,지각된 신뢰성,구매의도,샤오홍수

1. 서 론

샤오홍수는 커뮤니티 특성을 바탕으로 사용자들이 자신의 경험을 공유하고 이를 통해 다른 사용자들의 구매 결정을 돕는 소셜 커머스 생태계를 구축하고 있다. 특히 KOC(Key Opinion Consumer) 모델의 부상은 소비자들이 신뢰할 수 있는 실제 경험을 바탕으로 한 콘텐츠에 대한 신뢰로부터 출발하는데, 그 결과 KOC는 현재 브랜드 홍보와 판매에 중요한 역할을 하고 있다. KOC의 진정성과 평등한 상호작용 등의 특성은 소비자의 신뢰를 형성하고 구매 결정을 유도하는 데 있어서 점점 더 중요한 영향을 미치고 있다. 2023 KOL 마케팅 백서에

따르면, 샤오홍수 플랫폼에서 KOC의 비율은 81%에 이르며 상호작용이 활발한 게시물 중 초급 사용자들이 34%의 콘텐츠에 기여하고 있다고 한다(묘진 시스템과 전국 광고 연합회, 2023).

2. 이론적 배경

소셜 미디어 플랫폼의 활성화로 브랜드 정보가 빠르게 확산되었지만, 소비자는 정보의 진위를 구분하기 어려워졌다. Goh 등의 연구에 따르면, 허위 정보가 브랜드 신뢰를 손상시킬 수 있으며, 이를 해결하기 위해 브랜드 마케터들

은 인플루언서를 활용하기 시작했다. Woods(2016)는 소셜 미디어 인플루언서가 전통적 마케팅보다 소비자의 신뢰를 더 쉽게 얻을 수 있다고 주장하며, 이를 인플루언서 마케팅이라 부른다.

KOL(오피니언 리더)과 KOC(키 오피니언 소비자)는 소셜 미디어 인플루언서로서, KOC는 KOL보다 영향력은 적지만 더 진정성 있는 콘텐츠를 제공하여 신뢰를 구축한다. KOC 개념은 Yue-Long Yan이 처음 제안한 것으로, 진정성과 신뢰가 KOC의 주요 특성이다.

국제 연구에서는 KOC를 마이크로 인플루언서와 비교하며, Bijen(2017)은 마이크로 인플루언서가 KOL보다 높은 브랜드 신뢰와 구매 의도를 얻을 수 있다고 강조했다. 또한, 소비자들은 마이크로 인플루언서를 일반 소비자와 더 가깝게 느끼며, 그들의 의견은 개인 경험에서 비롯된 것으로 신뢰성이 높다고 한다.

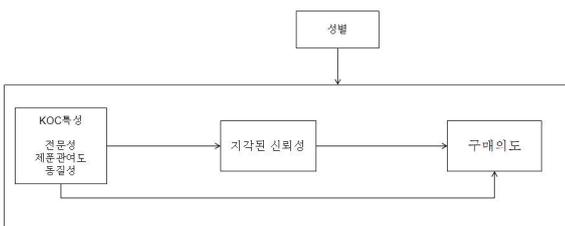
결론적으로, KOC와 마이크로 인플루언서는 진정성에 기반한 신뢰 매커니즘을 통해 브랜드 인식과 구매 결정에 영향을 미치며, 이는 소셜 미디어 마케팅에서 중요한 역할을 한다.

3. 연구 설계

3.1 연구 모델

본 모델에서 KOC의 세 가지 특성(전문성, 제품 관여도, 동질성)은 독립 변수로, 지각된 신뢰는 매개 변수로, 구매 의도는 종속 변수로 설정된다. 또한, 성별은 조절 변수로 작용하여 소비자가 KOC 특성에 대한 신뢰 수준에 영향을 미치고, 이는 소비자의 구매 결정에 추가적인 영향을 미친다.

구체적인 개념 모델은 다음과 같다:



3.2 연구 가설

H1: KOC의 특성의 지각된 신뢰성에 유의미한 정(+)의 영향을 미칠 것이다.

H1-1:전문성은 지각된 신뢰성에 유의미한 정(+)의 영향을 미칠 것이다.

H1-2:제품 관여도는 지각된 신뢰성에 유의미한 정(+)의 영향을 미칠 것이다.

H1-3:동질성은 지각된 신뢰성에 유의미한 정(+)의 영향을 미칠 것이다.

H2:지각된 신뢰성은 구매의도에 유의미한 정(+)의 영향을 미칠 것이다.

H3:KOC의 특성의 구매의도에 유의미한 정(+)의 영향을

미칠 것이다.

H3-1:전문성은 구매의도에 유의미한 정(+)의 영향을 미칠 것이다.

H3-2:제품관여도는 구매의도에 유의미한 정(+)의 영향을 미칠 것이다.

H3-3:동질성은 구매의도에 유의미한 정(+)의 영향을 미칠 것이다.

H4:지각된 신뢰성은 KOC의 특성과 구매의도 간의 관계를 매개할 것이다.

H4-1:지각된 신뢰성은 전문성과 구매의도 간의 관계를 매개할 것이다.

H4-2:지각된 신뢰성은 제품 관여도와 구매의도 간의 관계를 매개할 것이다.

H4-3:지각된 신뢰성은 동질성과 구매의도 간의 관계를 매개할 것이다.

H5: 성별이 본 연구의 개념적 모델에서 각 가설 간의 관계를 조절할 것이라는 가정을 제시한다.

4. 결 론

연구 결과에 따르면, KOC 특성이 지각된 신뢰 및 구매 의도에 미치는 영향은 성별에 따라 조절 효과가 있는 것으로 나타났다. 동질성은 여성 집단에서 지각된 신뢰에 더 큰 영향을 미쳤으며, 이는 여성 소비자가 KOC와의 유사성을 바탕으로 신뢰를 형성하는 경향이 크다는 것을 의미한다. 또한, 전문성은 여성 집단에서 구매 의도에 더 큰 영향을 미쳤으며, 이는 여성 소비자가 구매 결정을 내릴 때 KOC의 전문 능력을 더 중시한다는 것을 보여준다. 반면 남성은 KOC와의 유사성을 더 중요하게 여긴다.

따라서, 브랜드는 성별에 따른 차이를 고려해 KOC를 선정할 필요가 있다. 남성 소비자의 경우, 목표 소비자와 유사성을 지닌 KOC를 선택함으로써 구매 의도를 효과적으로 높일 수 있다. 여성 소비자를 대상으로 할 때는, 동질성뿐만 아니라 특정 제품 분야에서 높은 전문성을 가진 KOC를 선택하여, 다차원적인 신뢰를 통해 그들의 구매 결정을 촉진할 수 있다.

참 고 문 헌

[1] 김우빈·추호정(2019). SNS 패션 인플루언서 진정성이 팔로워 행동 의도에 미치는 영향: 팬십의 매개효과를 중심으로. 한국외국어학회지, 43(1), 17-32.
 [2] 김상조·정선미(2014), SNS에서의 외식정보정달자특성과 정보가치가 이용자의 외식업체 방문의도에 미치는 영향, 인터넷전자상거래연구, 14(6), 127-146.
 [3] 김재훈, 백림정, 한진수(2021). SNS 관광 인플루언서의 전문성과 매력성이 방문의도에 미치는 영향: 고객 태도의 매개효과를 중심으로. 호텔경영학연구, 30(1), 103-115.

은유적 내러티브를 활용한 ESG 환경 영상 디자인

문성예, 한지에¹⁾

조선대학교 디자인공학과

e-mail : msy283@naver.com, jijilion@chosun.ac.kr

Video Design for ESG focused on Environment applied on Metaphorical Narratives

Seong-Ye Moon, Ji-Ae Han

School of Informatics & Product Design, Chosun Univ.

요 약

본 연구는 소비자가 환경 보호의 진정한 의미와 중요성을 재고하며, 기업과 기관이 진정성 있게 ESG 원칙을 실천할 것을 촉구하는 설득영상 “Marionetter”를 디자인하고 그 의미를 제시했다. ESG 개념과 환경에 관한 고찰을 통해 메시지를 도출하고, 은유, 스토리텔링, 내러티브 요소를 구성했다. 피노키오 마리오네트 인형과 세트장으로 그린위싱의 주체와 환경을 은유적으로 표현했으며, 그린위싱의 사례들을 네 개의 에피소드로 구성해 스토리텔링과 내러티브 요소들로 표현했다.

1. 서 론

최근 기후변화 등으로 인해 많은 사람이 환경에 주목하게 되면서, 기업이 사회에 미치는 영향력과 비재무적 지표가 기업의 실질적 가치 평가에서 더 중요하다는 인식이 확산하고 있다. 그러나 많은 기업이 ESG(Environmental, Social, Governance) 경영을 내세우면서도 그 의미와 방향을 정확히 이해하지 못하는 경우가 많고, 이를 악용하는 경우가 늘어나고 있다. 그중에서, 친환경 효과를 과장하거나 환경 친화 이미지를 조작하는 ‘그린워싱’의 사례가 증가하고, 그에 따른 문제는 점점 더 심각해지고 있다. 이러한 환경적·사회적 문제점에 대한 경각심을 알리기 위해 환경에 초점을 두고 그 중요성을 알리는 영상을 디자인하고자 한다. 메타포와 스토리텔링을 사용하는 것은 정보 수용자에게 복잡한 주제를 쉽게 이해시키고, 감정적 연결을 강화하며, 중요한 메시지를 오래 기억하게 하는 효과적인 방법이기 때문에, 이를 활용해 영상 디자인을 진행하였다.

2. 영상 메시지 도출을 위한 ESG 고찰

‘ESG’ 개념은 2004년 4월 유엔이 “Who Cares Wins”라는 유엔책임 투자원칙(UNPRI)을 발표하면서 ESG 기준을 투자 결정 과정에 통합할 것을 제안한 것에서 시작됐다. 이는 기후변화와 팬데믹으로 인한 심각한 환경 문제로 인해 빠른 속도로 공감대를 형성하게 되었고, 환경적 요소·사회적 책임과 지배구조를 고려한 경영전략 즉, ESG 경영에 영향을 끼쳤다.

기업은 경영 과정에서 탄소배출 감소 등과 같이 환경에 미치는 영향을 고려해야 하며, 사회적 책임을 다해야 하며, 투명하고 윤리적 경영을 하기 위한 지배구조를 형성해

야 한다. 환경 측면에서는 탄소배출 감축과 에너지 효율성 향상, 재생 가능 에너지 사용 등을 통해 기후변화에 대응하며, 자원 관리, 폐기물 처리, 오염 방지 등의 활동으로 환경적 책임을 다해야 한다. 사회 측면에서는 노동자 권리 보호, 인권 존중, 다양성 포용을 통해 긍정적인 기업 문화를 조성하고, 지역사회 기여와 공정 거래를 통해 외부의 신뢰를 확보해야 한다. 지배구조 측면에서는 투명한 의사결정과 책임 경영을 강조하며, 특히 그린위싱을 방지하고 이사회 구성의 독립성과 윤리적 경영을 통해 기업의 신뢰성과 지속 가능성을 강화해야 한다.

표1은 ESG의 개념과 환경 관련된 키워드를 정리한 것이다. ESG의 세 가지 축인 환경, 사회, 그리고 지배구조를 종합적으로 고려할 때, ESG 환경 영상은 단순한 정보제공을 넘어, 정보 수용자가 환경 문제에 대한 비판적 인식을 하도록 유도하는 역할을 해야 한다. 이를 위해 영상 메시지는 환경에 미치는 영향을 강조하면서 기업의 책임을 명확히 제시함으로써, 정보 수용자들에게 지속 가능한 미래를 위한 실질적인 행동을 촉구해야 한다.

(표 1) ESG의 환경 관련 키워드 도출

구분	개념	환경 관련 키워드
Environmental	환경에 미치는 영향	온실가스 배출 감소, 탄소 배출 감소
Social	사회적 책임 수행	환경 정의, 기후 평등, 지속가능한 커뮤니티, 생태계 복원
Governance	경영의 투명성과 윤리성	그린워싱, 그린 라이팅, 그린 라벨링, 그린 시프팅

3. 설득형 메시지 전달을 위한 은유적 내러티브

설득형 영상에서 은유, 스토리텔링, 내러티브는 메시지 전달과 청중의 감정적 참여를 극대화하는 데 필수적인 역

1) 교신저자

할과 기능을 수행한다. 하나의 개념을 다른 개념에 빗대어 표현하는 은유는 복잡하거나 추상적인 아이디어를 보다 구체적이고 이해하기 쉽게 만들어준다. 스토리텔링은 청중의 감정을 자극해 메시지를 보다 개인적이고 친밀하게 전달하는 공감의 수단이 된다. 내러티브는 영상의 전체 구조를 형성하는 요소로 스토리텔링의 큰 틀을 제공하며, 영상의 논리적 흐름과 일관성을 유지하면서 메시지를 단계적으로 이해하게 하는 역할을 한다.

표2는 킨드라 홀(Kindra Hall)의 저서 「스토리의 과학」과 내러티브 관련 문헌연구를 토대로 은유, 스토리텔링, 내러티브에서 중요한 요소와 설득형 메시지 영상에서의 역할을 정리한 것이다.

(표 2) 은유, 스토리텔링, 내러티브의 중요요소와 역할

구분	중요 요소	역할
은유	시각적 매개체	추상적 개념의 구체화
	분명한 캐릭터	공간, 몰입, 경험 전달
스토리텔링	진실한 감정	진정성 부여, 감정적 연결
	중요한 순간	클라이맥스, 전환점, 주제 명확화
	구체적 디테일	현실감 부여, 생생한 이미지 제공
내러티브	역사적 맥락	인물, 사건, 전체적 분위기 이해
	정서적 연결	중요한 감정적 요소 구성
	전략적 전환점	내러티브의 흐름 조절, 강력한 인상 형성

4. ESG 환경 영상 “Marionetter”의 은유적 내러티브

"Marionetter"(문성예, 이유안, 2024)는 KBS와 시청자미디어센터 주관 2024 캠퍼스 온에어-숨쉬는 지구에서 시청자미디어재단 이사장상을 수상한 작품으로, 그린워싱의 기만적 본질과 이를 주도하는 조작자들의 존재가 표현됐다. 나아가, 그린워싱이 단순한 기업의 마케팅 전략을 넘어서는 복잡한 조작과 기만의 구조임을 시사하며, 정보 수용자가 이에 대해 비판적으로 인식하고 대응할 필요성을 강조했다. 영상 제목인 "Marionetter"는 "마리오네트를 조종하는 사람"이라는 의미를 담고 있다. 그린워싱의 작동 원리와 그로 인한 사회적 영향력을 암시하는 강력한 상징적 의미를 내포하고 있다.

(표 3) Marionetter의 은유, 내러티브 구성요소와 표현

내러티브/스토리텔링 구성요소	Marionetter의 표현
인물/캐릭터	피노키오 마리오네트 인형, 코끼리, 달팽이, 고양이, 검은색 말하는 종이인형
배경/공간	마리오네트 인형극 세트장
사건/진실한 감정	거짓말을 할수록 길어지는 코, 의심 없이 믿고 사용했다 당면한 문제
중요한 순간	홀러내리는 검정색 잉크로 망가진 세트장, 실이 끊어지면서 떨어지는 피노키오
문구/카피/멘트	당신도 모르는 사이 거짓말을 할 수도, 당할 수도 있다.
BGM/음향	경각심을 알려주는 무거운 BGM
색채	서서히 낮아지는 채도, 경고표시의 빨간색



(그림 1) Marionetter의 구성

영상의 인물/캐릭터인 피노키오 마리오네트는 그린워싱을 조장하는 지배구조에 조정되는 기업들, 코끼리, 달팽이, 고양이는 거짓된 친환경 이미지를 내세우는 기업들로 인해 고통받는 소비자들의 상황을 비유적으로 표현한 것이다. 감정적 연결과 현실감을 부여하기 위해, 의심 없이 친환경 제품을 믿고 사용하다가 발생하는 사건을 세 개의 에피소드로 구성했으며, 각 에피소드에서 주제를 나타내기 위해 거짓말을 할수록 길어지는 피노키오의 코를 강조했다. 중요한 순간, 즉 전략적 전환점을 제공하기 위해 홀러내리는 검정 잉크로 망쳐진 세트장과 실이 끊어지면서 주저앉은 피노키오 인형으로 영상의 클라이맥스를 구성했다. 이를 통해 그린워싱이 단순히 일부 소비자만을 속이는 문제가 아니라, 광범위한 사회적영향을 미칠 수 있다는 점을 시사한다. 그린워싱이 만연할 경우, 진정한 환경 보호 노력은 왜곡되고, 사회 전체가 잘못된 방향으로 나아갈 위험이 있음을 표현하고 있다.

5. 결론

은유적 내러티브를 활용해 소비자가 환경 보호의 진정한 의미와 중요성을 재고하며, 기업과 기관이 진정성 있게 ESG 원칙을 실천할 것을 촉구하는 설득영상 “Marionetter”를 제안하였다. 이 작품에 사용된 은유적 요소와 내러티브 요소가 ESG 원칙에 대한 기업의 무지와 사회의 무관심이 결합되었을 때, 단순한 환경 문제를 넘어선 사회적, 경제적 파국으로 이어질 수 있음을 전달하는데 효과적이었다고 판단된다.

참고 문헌

- [1] 김은주, 김건. “칠성사이다 TV광고 내러티브의 구성요소와 공감반응”, 『한국디자인포럼』 제23권 제1호, 2018.
- [2] 김남두, 이창호. “정보 사유와 공유의 레퍼토리와 은유적 내러티브”, 『한국언론학회』 제49권 제6호, 2005.

산업환경에서 개인 보호 장비 감지를 위한 개선된 YOLOv8 모델 설계

박석우¹, 이영석², 나인호²

¹국립군산대학교 전자정보공학부, ²국립군산대학교 컴퓨터소프트웨어학부
e-mail : {p_sw, leeys, ihra}@kunsan.ac.kr

A Design of Enhanced YOLOv8 Model for Personal Protective Equipment Detection in Industrial Environments

Seok-Woo Park¹, Young-Seok Lee², In-Ho Ra²

¹School of Electronics and Information Engineering
Kunsan National University

²School of Computer and Software
Kunsan National University

요 약

산업 환경에서의 안전관리는 근로자의 생명과 건강을 보호하기 위한 필수요소로서 이를 효율적으로 관리하지 못할 경우 사고와 부상으로 인한 재정적 손실 및 기업의 경쟁력 저하로 이어질 수 있다. 이러한 상황에서 개인 보호 장비(PPE)의 착용 여부를 실시간으로 감지하는 시스템은 사고를 예방하는데 있어 중요한 역할을 수행한다. 그러나 이와 관련된 기존 연구들은 실제 산업 환경의 복잡한 조명 변화, 다양한 각도, 작업자의 움직임 등의 조건에서 성능 저하, 고사양 하드웨어와 복잡한 네트워크 구조로 인해 실시간 적용이 어려운 문제를 지닌다. 본 연구에서는 이러한 문제를 해결하기 위해 GhostConv, GhostBottleneck, C3Ghost, BiFPN, CBAM 등의 모듈을 조합하여 개선된 YOLOv8 모델을 제안한다. 특히 CBAM을 활용해 피쳐 맵에서 중요한 정보를 강조함으로써 PPE 감지 성능을 향상할 수 있으며, 다양한 크기와 위치의 보호 장비를 효과적으로 인식할 수 있다. 향후 연구에서는 제안된 모델의 성능을 다양한 산업 환경에서 정량적으로 평가하여 실질적인 현장 적용 가능성을 검증할 예정이다.

1. 서 론

산업 환경에서의 안전관리는 근로자의 생명과 건강을 보호하며, 특히 효율적인 안전관리는 사고와 부상 위험을 최소화할 수 있다. 안전 미흡으로 발생하는 사고는 의료비를 포함한 각종 보상, 생산성 저하 등으로 인해 막대한 재정적 손실을 초래하고 기업의 성장지속 가능성과 경쟁력에 부정적 영향을 미칠 수 있다[1]. 산업 현장에서 발생하는 안전사고는 추락, 기계 결함, 화재 등 다양한 원인으로 발생하며, 이를 예방하기 위한 안전관리의 중요성이 꾸준히 강조되고 있다. 이러한 측면에서 컴퓨터비전 기술은 산업 안전관리에 있어 혁신적인 역할을 수행하고 있으며, 특히 개인 보호 장비(PPE) 감지 시스템은 실시간 작업 환경 모니터링과 위험 요소의 식별을 통해 즉각적인 대응을 가능하게 한다[2, 3]. 그러나 기존 PPE 감지 시스템은 여전히 다음과 같은 여러 가지 한계를 지니고 있다.

첫째, 기존 연구는 표준 환경에서 수집된 데이터를 바탕으로 모델을 개발하여 실제 산업 환경의 복잡한 조명 변화, 다양한 각도, 작업자의 움직임 등의 조건에서 성능이 저하되는 문제가 있다[4]. 둘째, 고성능을 위해 복잡한 네트워크 구조와 고사양 하드웨어를 요구하는 경우가 많아 실시간 적용에 한계가 있어 산업현장에서 즉각적인 안전

관리에 어려움을 겪고 있다[5]. 셋째, 다양한 보호 장비의 동시다발적 인식에 어려움을 겪으며, 작업자가 착용한 보호 장비를 제대로 감지하지 못하는 문제가 있다[6].

이러한 문제를 해결하기 위해 본 연구에서는 산업 환경에서 개인 보호 장비 감지를 위한 새로운 모델을 제안한다. 제안된 모델에서는 GhostConv, GhostBottleneck, C3Ghost, Attention, BiFPN 등을 조합하여 PPE 감지 성능을 향상시키도록 한다. 여기서 GhostConv와 GhostBottleneck은 경량화된 연산을 통해 실시간 성능과 정확한 감지를 제공하며, BiFPN과 Attention 모듈은 다양한 크기와 위치의 보호 장비를 효과적으로 감지하도록 한다. SimC2f와 SPPF를 통해 피쳐 맵을 통합하고 강조하여 복잡한 산업 환경에서도 높은 성능을 발휘할 수 있다.

2. 모델 설계

본 연구에서는 산업 환경에서의 개인 보호 장비(PPE) 감지를 위해 개선된 YOLOv8 모델 구조를 제안한다. 제안된 모델은 효율적인 피쳐 추출, 다양한 크기의 객체 감지, 그리고 실시간 처리 성능을 목표로 GhostConv, GhostBottleneck, C3Ghost, BiFPN, CBAM 등을 조합하여 안전관리 모니터링 시스템의 성능을 최적화한다. 제안된 모델은 그림 1과 같이 Backbone, Neck, Head로 구성된다.

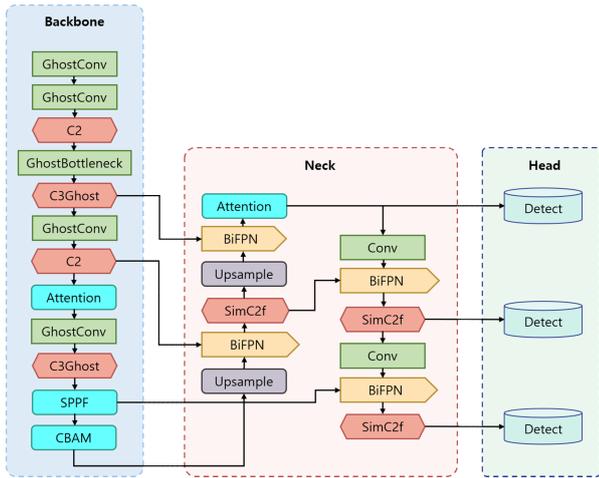


그림 1. 개선된 YOLOv8 모델 구조

2.1 Backbone

Backbone은 입력 이미지에서 피처를 추출하는 역할을 하며, 제안된 구조에서는 GhostConv와 GhostBottleneck을 사용해 경량화된 연산을 통해 효율적인 피처 추출을 가능하게 한다. GhostConv는 저비용의 실시간 처리에 적합하며, GhostBottleneck은 중요한 피처를 보존하면서 연산 효율을 높인다. C3Ghost는 피처 맵을 여러 경로로 분리하고 결합하여 정보 혼합과 깊이 최적화를 통해 정교한 피처 추출을 지원한다. 또한, Attention 모듈은 중요한 피처를 강조하여 PPE와 같은 작은 객체의 검출 성능을 향상시키며, SPPF와 CBAM은 다양한 크기의 피처를 통합하고 강조하여 객체 인식을 개선한다. 특히 CBAM은 채널 및 공간 주의 메커니즘을 결합해 중요한 채널과 위치를 강조하며, CBAM의 Channel 및 Spatial Attention은 다음 식(1)과 같다.

$$\begin{aligned} M_c(F) &= \sigma(W_1(ReLU(W_0(F^T)))) \\ M_s(F) &= \sigma(f^{7 \times 7}(F)) \end{aligned} \quad (1)$$

여기서 $f^{7 \times 7}$ 는 7×7 Convolution 연산이다.

2.2 Neck

Neck은 Backbone에서 추출된 피처를 결합하여 최종적으로 Head로 전달하기 전에 피처를 강화하는 역할을 한다. 본 모델에서는 BiFPN(Bidirectional Feature Pyramid Network)를 사용하여 다중 해상도에서 피처를 결합하고, 정보의 흐름을 최적화하여 객체 검출의 정확도를 향상시킨다. BiFPN은 상하 양방향으로 피처 맵을 전달하며, 중요한 피처 정보를 강조함으로써 더욱 정교한 감지를 가능하게 한다. BiFPN의 노드 간 연산은 다음 식(2)와 같다.

$$P_i = \frac{\sum_{j \in \text{input nodes}} \omega_j P_j}{\sum_{j \in \text{input nodes}} \omega_j + \epsilon} \mathbf{d} \quad (2)$$

Neck 구조에 포함된 SimC2f 모듈은 간소화된 C2f 구조로서 필요한 피처만을 효율적으로 추출하여 연산량을 줄이면 중요한 정보는 지속적으로 유지한다. 또한, Upsample 레이어는 피처 맵의 해상도를 높여 작은 크기의 PPE 감지 성능을 강화한다.

2.3 Head

Head는 최종적으로 객체를 검출하고 분류하는 단계로, Detect 모듈이 포함되어 있다. 이 모듈은 여러 레벨의 피쳐 맵을 사용하여 다양한 크기와 위치의 PPE를 정확히 인식할 수 있다. 특히, BiFPN에서 전달된 피쳐 맵을 활용하여 높은 정확도를 유지하며 실시간 감지가 가능하도록 최적화한다.

3. 결론

본 논문에서는 CBAM을 통해 피쳐 맵에서 중요한 정보를 강조하여 PPE 감지의 정확도를 크게 향상시킬 수 있으며, GhostConv와 BiFPN의 경량화된 구조를 통해 연산량 감소 및 실시간 처리 성능을 유지할 수 있는 모델을 제안하였다. 또한, 조명 변화, 복잡한 배경, 다양한 각도 등 실제 산업 환경에서도 높은 성능을 발휘할 수 있도록 하였다. 향후 연구에서는 제안된 YOLOv8 모델의 성능을 분석하여 각 모듈의 효과를 정량적으로 평가하고, 다양한 산업 환경에서의 적용 가능성을 검증할 예정이다. 마지막으로, 다양한 조건에서 모델의 성능을 평가하여 범용성을 높이고 실질적인 현장 적용 가능성을 검증할 계획이다.

Acknowledgement

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2023RIS-008).

참고 문헌

- [1] Kim Dong Koo, Sun young Park. "An Analysis of the Effects of Occupational Accidents on Corporate Management Performance", Safety Science, Vol. 138, 2021.
- [2] A. Hayat, M. D. Fernando. "Deep Learning-based Automatic Safety Helmet Detection System for Construction Safety", Applied Sciences, Vol. 12, No. 16, 2022.
- [3] M. Abouelyazid, "YOLOv4-based Deep Learning Approach for Personal Protective Equipment Detection", Journal of Sustainable Urban Futures, Vol. 12, Vol. 3, 2022.
- [4] R. Mukherjee, M. Bessa, P. Melo-Pinto, A. Chalmers, "Object Detection under Challenging Lighting Conditions using High Dynamic Range Imagery", IEEE Access, Vol. 9, pp. 77771-77783, 2021.
- [5] G. Gallo, F. Di Rienzo, F. Garzelli, P. Ducange, C. Vallati, "A Smart System for Personal Protective Equipment Detection in Industrial Environments based on Deep Learning at the Edge", IEEE Access, Vol. 10, pp. 110862-110878, 2022.
- [6] C. Shi, D. Zhu, J. Shen, Y. Zheng, C. Zhou, "GBSG-YOLOv8n: A Model for Enhanced Personal Protective Equipment Detection in Industrial Environments", Electronics, Vol. 12, No. 22, 2023.

초기 및 후기 융합 기법을 활용한 개선된 MDVC 모델

조성민, 최경주
 충북대학교 소프트웨어학과
 e-mail : csmcc2151@gmail.com, kjcheoi@chungbuk.ac.kr

An Improved MDVC Model Utilizing Early and Late Fusion Techniques

Seongmin Cho, Kyung Joo Cheoi
 Dept. of Computer Science, Chungbuk National University

요약

본 논문에서는 기존 비디오 캡션 생성 모델인 MDVC(Multi-modal dense video captioning)의 구조를 변경하여 시각 및 오디오 특징은 초기 융합하고 음성 특징은 후기 융합하는 방식으로 새로운 구조의 비디오캡셔닝 모델을 제안하였다. 실험 결과, 제안된 모델은 성능평가에서 기존 모델보다 전반적으로 향상된 결과를 보였으며, 특히 Learned Proposals에서 높은 성능을 보였다. 이러한 연구 결과는 멀티모달 데이터를 융합할 때, 후기 융합 방식을 단독으로 사용하는 것보다 초기 융합과 후기 융합을 모두 활용하는 방식이 더욱 효과적임을 입증한다.

1. 서론

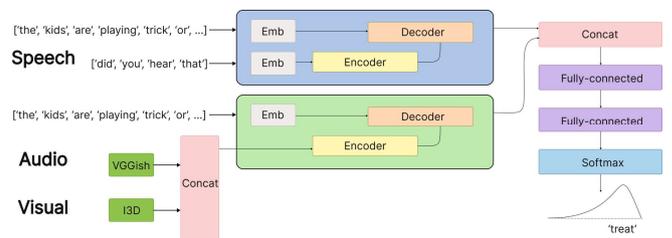
기존의 비디오 캡션 생성 모델들은 주로 시각 정보에만 의존하여 개발되었다. 하지만 시각 정보만으로는 비디오에서 발생하는 모든 사건의 맥락을 완전히 이해하기 어렵고, 비디오 이해에 도움이 되는 오디오 및 음성 신호가 무시될 수 있다. 따라서 최근 연구 동향은 비디오의 상황을 정확히 이해하기 위하여 비디오 정보뿐만 아니라 오디오 정보 등을 함께 고려하는 멀티모달(Multi-modal)을 사용한다 [1-6]. 본 논문에서는 기존의 MDVC(Multi-modal dense video captioning)[3]를 개선한 새로운 모델을 제안한다. 기존의 모델은 오디오, 시각(visual), 음성 신호를 사용한 멀티모달 시스템이지만, 각각의 특징이 추출되면, 각각의 특징이 개별 트랜스포머(Transformer) 모델에 독립적으로 입력되어 처리된 후, 처리된 개별 특징들을 융합하는 후기 융합 방식을 채택하였다. 이렇게 특징이 개별적으로 처리된 후 마지막에 융합되는 후기 융합 방식은 각 모달리티의 정보를 별도로 처리함으로써, 시각 정보와 오디오 정보 간의 상호 보완적 관계를 충분히 반영하지 못하는 한계가 있다. 본 논문에서는 이러한 한계를 보완하기 위하여, 초기 융합과 후기 융합을 모두 사용하여 각 모달리티 간의 상호작용을 극대화함으로써 기존 모델 대비 성능이 개선된 모델을 제안하였다.

2. 제안하는 모델

2.1 제안하는 모델의 구조도

그림 1은 제안된 시스템의 전체적인 구조를 보여준다. 먼저 오디오 정보는 VGG 네트워크 구조를 기반으로 한 오디오 특성 추출 모델인 VGGish를, 시각(visual) 정보는

I3D(Inflated 3D) 모델을 사용하여 특징이 추출되고, 이렇게 추출된 2가지 종류의 특징이 융합(초기 융합)된 후, 트랜스포머의 입력으로 사용된다. 음성 신호는 2개의 특징과 별개로 독립적인 트랜스포머의 입력으로 사용된다. 음성 신호는 유튜브(YouTube) 비디오에서 제공하는 CC(Closed Captions)로 사람 간의 대화나 말소리를 의미한다. 오디오는 비디오에서 발생하는 모든 소리를 의미하며, 여기에는 배경 소음, 음악 등이 포함된다. 결국 음성 특징과 초기에 미리 융합된 시각-오디오 특징은 개별적으로 트랜스포머에 의해 처리된 후, 다시 하나로 융합(후기 융합)되어 최종 결과를 도출하게 된다. 또한, 트랜스포머 훈련 과정에서 정확한 캡션 예측을 위해 교사 강제(teacher forcing) 기법을 적용하였다. 이 기법으로 이전 시점에서 디코더가 생성한 단어를 임베딩 과정을 거쳐 벡터화한 후, 디코더의 입력으로 사용하였다.



(그림 1) 제안된 모델의 전체적인 구조도

2.2 초기 융합과 후기 융합

초기 융합은 서로 다른 모달리티에서 추출된 특징을 결합하여 하나의 벡터로 만들고, 이를 기반으로 학습을 진행하는 방식이다. 본 논문에서는 입력 단계에서 오디오와 시

각 모달리티의 특징을 연결(Concatenate) 방법을 사용하여 하나의 통합된 입력 벡터로 처리한다. 후기 융합은 서로 다른 모달리티 별로 독립적인 트랜스포머 모델을 통해 학습한 후, 각 모달리티에서 도출된 결과를 결합하여 최종 결정을 내리는 방식이다. 본 논문에서는 모달리티 별로 독립적인 학습을 진행한 후, 예측 단계에서 각 모달리티의 결과를 연결(Concatenate) 방법으로 결합한다.

3. 실험 및 결과

3.1 데이터세트와 평가 지표

본 논문에서는 ActivityNet Captions 데이터세트[7]를 사용하여 실험을 수행하였다. 데이터세트는 약 20k개의 YouTube 비디오로 구성되어 있으며 훈련용 비디오(10k개), 검증용 비디오(5k개), 테스트용 비디오(5k개)로 분할되어 있다. 각 비디오는 평균 3.65개의 문장을 포함하고 있으며 문장 당 평균 길이는 13.48단어로, 평균적으로 36초 길이로 되어 있다. 또한, 음성 신호는 YouTube ASR(Automatic Speech Recognition) 시스템에 의해 CC(Closed Captions)로 변환되었고, 이를 음성 신호로 사용하였다. 성능평가 지표로는 기존 연구들에서 많이 사용된 BLEU(B@3, B@4)와 METEOR를 사용하였다. BLEU(Bilingual Evaluation Understudy Score)는 생성된 텍스트와 정답 텍스트 간의 n-그램 일치율을 측정하여 점수를 산출하며, 점수가 높을수록 모델이 정답 텍스트와 유사한 텍스트를 생성했음을 의미한다. METEOR (Metric for Evaluation of Translation with Explicit ORdering)는 어휘적 변형, 동의어, 어순 등을 고려하여 텍스트 간의 유사성을 평가하며, 점수가 높을수록 모델 성능이 우수함을 의미한다.

3.2 실험 결과

제안된 모델의 성능 평가는 기존 연구 중 성능이 비교적 높은 몇 개의 연구들[1,3]과 비교하였다. 제안된 모델인 ‘Ours, no missings’는 데이터에서 누락된 정보를 제거한 후 학습된 모델로, GT Proposals(사람이 작성한 정답 캡션)에서는 일부 높은 성능이 나왔고 Learned Proposals(모델이 생성한 정답 캡션)에서는 전반적으로 모두 높은 성능이 나왔다.

4. 결론

본 논문에서는 기존의 MDVC 모델의 구조를 후기 융합 방식만 단일적으로 사용하는 것이 아닌 초기 융합과 후기 융합 방식을 모두 활용하는 방식으로 변경한 새로운 구조를 제안하였다. 융합 방식의 변경을 통해 다양한 모달리티 간의 상호작용을 강화하고자 하였다. ActivityNet Captions 데이터세트를 사용하여 실험한 결과, 제안된 모델은 성능평가에서 전반적으로 향상된 결과를 보였다. 이는 멀티모달 결합에서 후기 융합보다 초기 융합과 후기 융합을 모두 활용하는 방식이 더욱 효과적인 방법임을 입

증한다. 다만, 제안하는 모델은 GT Proposals에서 일부 성능이 저하된 부분이 있었다. 향후 연구에서는 이러한 일부 성능 저하를 극복하기 위해 각 모달리티의 중요성을 고려한 가중치 조정을 통해 보다 정교한 융합 방식을 연구하고자 한다.

(표 1) 기존 연구들과의 성능 비교

Method	GT Proposals			Learned Proposals		
	B@3	B@4	Meteor	B@3	B@4	Meteor
Rahman et al.[1]	3.04	1.46	7.23	1.85	0.90	4.93
MDVC[3]	4.12	1.81	10.09	2.31	0.92	6.80
Ours	4.45	1.93	9.98	2.41	0.99	6.86
MDVC, no missings[3]	5.83	2.86	11.72	2.60	1.07	7.31
Ours, no missing	5.92	2.83	11.46	2.68	1.14	7.35

Acknowledgment

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2019-0-01183)의 지원을 받아 작성됨.

참고문헌

- [1] Tanzila Rahman, Bicheng Xu, and Leonid Sigal. “Watch, listen and tell: Multi-modal weakly supervised dense event captioning”, In ICCV, 2019. pp.1-6
- [2] Jack Hessel, Bo Pang, Zhenhai Zhu, and Radu Soricut. “A case study on combining ASR and visual features for generating instructional video captions”, In CoNLL, 2019. pp.1-5
- [3] Vladimir Iashin and Esa Rahtu. “Multi-modal dense video captioning”, In CVPR, 2020. pp.1-6
- [4] Luo H, Ji L, Shi B, Huang H, Duan N, Li T, Li J, Bharti T, Zhou M. “UniVL: A Unified Video and Language Pre-Training Model for Multimodal Understanding and Generation”, In CVPR, 2020 pp.1-7
- [5] Botian Shi, Lei Ji, Yaobo Liang, Nan Duan, Peng Chen, Zhendong Niu, and Ming Zhou. “Dense procedure captioning in narrated instructional videos.”, In ACL, 2019. pp.6382-6386
- [6] Luo H, Ji L, Shi B, Huang H, Duan N, Li T, Li J, Bharti T, Zhou M. “UniVL: A Unified Video and Language Pre-Training Model for Multimodal Understanding and Generation”, In CVPR, 2020 pp.1-7
- [7] Ranjay Krishna, Kenji Hata, Frederic Ren, Li Fei-Fei, and Juan Carlos Niebles. “Dense-captioning events in videos”, In ICCV, 2017. pp.5

클라우드 관점에서의 스마트리터시에 관한 역량설계 연구

장계훈
중앙대학교 보안대학원 보안공학과 석사과정
e-mail : jkh1942@gmail.com

Study on competency design for Smart Literacy from a Cloud Perspective

Gaehoon Chang
Master's program in Security Engineering, Graduate School of Security,
Chung-Ang University

요 약

본 연구는 다양한 디지털 전환 과정에서 클라우드 컴퓨팅의 이해와 역량개발 방법에 대해 논의한다. 현대 사회에서 리터러시는 단순한 문자 해독 능력을 넘어, 디지털 미디어가 보편화됨에 따라 가상 환경을 이해하고 인식하는 능력으로 확장되었다. 이와 함께 특정 분야에서 요구되는 역량과 지식으로 세분화된 다양한 리터러시 개념이 등장하고 있다. 전통적인 IT 환경은 물리적 데이터 센터를 기반으로 하였으나, 현재는 클라우드 컴퓨팅을 통해 가상화된 환경으로 급격히 전환되고 있다. 이러한 변화 속에서 전통적인 IT 엔지니어와 최신 클라우드 컴퓨팅 기술을 보유한 엔지니어 간의 역량 차이는 명확하게 드러나며, 이들 기술은 대립적이기보다는 상호 보완적인 관계에 있다. 따라서, 두 그룹 간의 역량 격차를 줄이기 위한 프로그램 개발이 필요하다. 본 역량 설계 모형은 어떻게 프로그램을 개발할지에 절차를 마련하여 다양한 클라우드 컴퓨팅 환경에 적극적으로 대응하는 발판을 마련하고자 한다.

This study discusses the understanding and capability development methods of cloud computing in various digital transformation processes. In modern society, literacy has expanded beyond simple reading skills to include the ability to understand and recognize virtual environments, as digital media has become ubiquitous. Along with this, various concepts of literacy, segmented by the competencies and knowledge required in specific fields, have emerged. Traditional IT environments were based on physical data centers, but there is now a rapid transition to virtualized environments through cloud computing. In this changing landscape, the competency gap between traditional IT engineers and those equipped with the latest cloud computing technologies is clearly evident, and these technologies are in a complementary rather than adversarial relationship. Therefore, there is a need to develop programs aimed at reducing the competency gap between the two groups. From this perspective, we aim to establish procedures and flowcharts for program development to create a foundation for actively responding to various cloud computing environments.

효율적인 해킹공격 대응을 위한 LLM 기반 허니팟 구축 아키텍처 연구

이진욱
중앙대학교 보안대학원 보안공학과
e-mail : leejw79@cau.ac.kr

Research on LLM-based honeypot construction architecture for efficient response to hacking attacks

Jinwook Lee
Department of Security Engineering, Graduate School of Security,
Chung-Ang University

요 약

LLM은 대규모 데이터를 활용하여 학습된 자연어 처리 모델로서 인간이 사용하는 자연어를 이해 할 수 있어 산업 여러 분야에 활용되고 있는 언어모델이다. 대표적으로 OpenAI의 GPT, Google의 BERT, Anthropic의 Claude와 같은 모델들이 있으며, 이런 모델들을 활용한 챗봇 서비스 등 다양한 생성형 AI 서비스가 개발되고 있다. 사이버보안 분야에서도 생성형 AI 기반 챗봇을 활용하여 보안솔루션의 정/오답 판단에 활용 가능한 정보를 제공하는 등 보조적인 기능을 제공하는 등 AI를 적극 활용하고 있다. 뿐만 아니라 최근에는 생성형 AI를 활용한 Honeypot에 대한 연구도 진행 되고 있어 실제 해커의 공격 Payload에 대응하여 생성형 AI가 Web/Was Server의 응답 패킷이나 OS의 명령어 실행 결과를 상황에 맞게 생성하여 제공하는 기능을 연구하고 있다. 본 연구에서는 WAF, Proxy 서버를 활용하여 정상 요청과 비정상(공격) 요청을 구분하여 정상 사용자의 요청일 경우 정상 서비스 경로를 통해 응답을 제공하고 해커와 같은 공격자의 요청일 경우 생성형 AI 기반의 Honeypot으로 트래픽을 돌려 공격 Payload에 맞는 콘텐츠를 생성하여 제공 함으로써 공격자를 격리시키고, 공격패턴을 수집하는 아키텍처를 제시한다. 본 아키텍처를 활용할 경우 서비스의 가용성을 확보하고 해커에게는 공격이 성공한 듯한 상황을 제공 함으로써 해커의 리소스를 소모시켜 서비스의 보안성을 더욱 향상할 수 있다. 또한 수집한 해커의 공격 Payload를 WAF, IPS와 같은 보안 솔루션의 정책에 반영하는 프로세스를 통해 더욱 강건한 서비스 아키텍처를 구성할 수 있다.

보안관제 시스템에서 방화벽 로그 분석을 위한 AI 적용 방안

김지훈
중앙대학교 보안대학원 보안공학과
e-mail : qtkamiru@cau.ac.kr

Application of AI for Firewall Log Analysis in Security Monitoring Systems

Jihun Kim
Department of Security Engineering, Graduate School of Security,
Chung-Ang University

요 약

보안 위협이 지속적으로 진화하고 증가하고 있어, 이를 탐지하기 위한 보안관제 시스템에서 수집하는 로그도 시간이 지남에 따라 점차 증가하고 있다. 특히, 방화벽의 로그는 솔루션 특성상 로그 양이 더욱 많아 보안관제에서는 특정 조건에 해당하는 것만 보거나 사후추적용으로 활용되는 경우가 많다. 본 연구는 AI 기술을 활용하여 방화벽 로그에서 비정상 로그를 탐지 및 예측할 수 있도록 하여 실시간 위협 탐지 성능을 개선하는 방안을 제안한다. 이 연구는 보안관제 시스템의 성능 향상과 더불어 빠르게 위협을 찾아내어 침해 피해를 줄이는데 기여할 수 있다.

제1금융권과의 비교를 통한 제2금융권의 적정 보안 체계 수립 및 실무적 보안 강화 방안: 중소 금융기관을 위한 효과적 대응 전략

이은지
중앙대학교 보안대학원
e-mail : ejlee9@cau.ac.kr

The Establishment of Appropriate Security Systems and Practical Security Reinforcement in Secondary Financial Institutions Through a Comparison with Primary Financial Institutions: Effective Response Strategies for Small and Medium-Sized Financial Institutions

Lee Eun Ji
Graduate School of Security at Chung-Ang University

요 약

금융 산업의 디지털화가 가속화되면서 사이버 보안의 중요성이 더욱 부각되고 있다. 특히 제2금융권에 속하는 중소 금융기관들은 제1금융권 대비 상대적으로 보안 체계가 미흡한 실정이다. 본 연구는 제1금융권과 제2금융권 간의 보안 수준 차이를 비교하고, 제2금융권에 적합한 보안 체계 수립과 실무적 보안 강화 방안을 제시하는 것을 목표로 한다. 규모 및 운영 범위, 규제 준수, 지배구조의 차이 등으로 인해 발생하는 보안 격차를 해결하기 위한 실질적인 방안을 제시하고, 이를 통해 중소 금융기관의 사이버 위협 대응 능력을 강화하는 데 기여하고자 한다.

호텔 내부자의 악의적 행위 리스크 평가 : 심리적 계약 위반 이론 중심으로

권중훈

중앙대학교 보안대학원 산업보안정책학과
e-mail : kbaio@cau.ac.kr

Risk assessment of malicious behavior by hotel insiders: Focusing on psychological contract violation theory

Joung Hun Gwon

Department of Technology Protection, Chung-Ang University

요 약

이 연구는 호텔 산업에서 내부자의 악의적인 행위 리스크를 심리적 계약 위반 이론을 중심으로 평가하는 것을 목적으로 한다. 내부자의 악의적인 행동은 호텔 운영과 평판에 매우 큰 피해를 줄 수 있으며, 이러한 문제는 회사와 고용자간 심리적인 계약이 깨질 때 발생한다. 본 연구는 심리적인 계약 위반이 내부자의 위협에 있어 매우 중요한 요인임을 강조하며 직원이 느끼는 배신감과 불만 등이 악의적인 행동으로 이어질 수 있음을 탐구한다. 주요 연구 내용은 심리적 계약위반과 내부자 위협 간의 관계, 위반의 구체적인 유형 등을 다루며 이러한 계약 위반이 내부자의 악의적 행위에 미치는 영향을 조사하기 위해 설문조사를 활용하여 양적 데이터를 수집한다. 연구결과를 바탕으로 호텔이 내부자 위협에 대한 심리 요인을 이해하고 이를 통해 호텔의 보안과 신뢰성을 강화하는데 기여하고자 한다.

AI 기술을 이용한 보안인증제도 활성화 방안

안진우
중앙대학교 보안대학원 보안공학과
e-mail : anjin227@cau.ac.kr

The Activation Plan of Security Certification System using AI

Jinwoo An
Department of Security Engineering, Graduate School of Security,
Chung-Ang University

요 약

정보보호 및 개인정보보호 관리체계 인증(ISMS, ISMS-P)은 기업의 정보보호 및 개인정보보호 관리체계가 적절하게 운영되고 있는지를 심사하고 인증하는 제도이다. 그러나 2013년 ISMS 인증이 의무화된 이후, 의무대상 기업은 관리체계 운영과 인증 과정에서 부담이 증가하였고, 인증을 받은 후에도 보안 사고가 발생하는 사례가 있었다. 이러한 문제는 짧은 시간 안에 기업의 전체 보안 현황을 파악하고 점검하는 것이 현실적으로 어려울 뿐만 아니라, 심사원 개인의 노하우나 역량에 따라 점검의 깊이나 관점이 달라져 결함을 놓치는 경우가 생기기 때문이다. 대규모 언어 모델(LLM)은 대량의 문서를 학습한 내용을 바탕으로 검색을 하거나 자연어 기반의 질의응답 기능을 제공한다. 본 연구에서는 이러한 LLM 기술을 활용해 인증심사 '사전 점검' 단계에서 기업의 보안 수준과 잠재적 결함을 미리 확인하고, 개선 방안을 조언받아 조치하는 방안을 제안한다. 이를 통해 기업의 정보보안 정책 문서나 IT 인프라의 보안 설정이 관련 법규 및 기업 정책과 일관되게 준수되고 있는지 점검하여, 인증 심사 절차의 효율성을 높이고 기업 전반의 보안 수준을 향상시킬 수 있다.

불만에서 혁신으로: 소비자의 부정적 리뷰 분석 기반 전기차 시장 확대 방안

왕문준, 최재원
순천향대학교 경영학과, 순천향대학교 경영학과
e-mail : dwzhzsa22@sch.ac.kr, jaewonchoi@sch.ac.kr

From Complaint to Innovation: How to Drive EV Market Development Through Negative Consumer Reviews

Wenjun Wang, Jaewon Choi
Soonchunhyang University, Department of Business Administration,
Soonchunhyang University, Department of Business Administration

요 약

전기차 시장이 확대됨에 따라 소비자의 기대도 커지고 있지만, 최근 일부 지역에서 자발적 화재 사고 등 안전 문제로 인해 전기차의 지하주차장 출입이 제한되는 사례가 발생하고 있다. 이러한 상황은 소비자의 불안과 불만으로 이어져 전기차에 대한 신뢰와 시장 성장에 부정적인 영향을 미칠 수 있다. 이 연구는 전기차에 대한 부정적인 소비자 리뷰를 분석하고 이러한 불만이 전기차 시장의 혁신과 성장을 촉진할 수 있는 잠재력을 탐구한다.

텍스트 마이닝 기법을 사용하여 온라인 리뷰 데이터를 분석한 결과, 배터리 성능, 충전 인프라 부족, 안전 관련 문제가 소비자들의 주요 불만 사항으로 나타난다. 이 연구는 기술 혁신, 서비스 개선, 정책적 대응을 통해 해결할 수 있는 과제를 제시한다. 특히 배터리 안전성 강화, 충전 인프라 확충, 화재 방지 시스템 도입 등 기술적 접근뿐만 아니라 정책적 지원 확대와 서비스 개선이 전기차 시장의 지속 가능성을 높이는 데 중요한 역할을 할 수 있다고 제안한다.

이 연구는 소비자 불만족이 전기차 산업의 장기적인 발전을 위한 혁신 동력이 될 수 있음을 강조하고 부정적인 피드백을 통해 더 나은 제품과 서비스를 개발할 수 있는 방법을 제시한다. 이처럼 소비자 불만족에 대한 학술적 연구는 전기차 산업 성장의 동력이 될 수 있다.

서사적 매력과 광고 전파력: 서사 수준이 바이럴 효과에 미치는 영향

유몽, 최재원

순천향대학교 경영학과, 순천향대학교 경영학과
e-mail : menliu224@sch.ac.kr, jaewonchoi@sch.ac.kr

The Narrative Appeal and Advertising Reach: The Impact of Narrative Levels on Viral Effects in Social Media

Meng Liu, Jaewon Choi

Soonchunhyang University, Department of Business Administration,
Soonchunhyang University, Department of Business Administration

요약

소셜 미디어의 부상으로 광고 서사 방식은 바이럴 전파에서 중요한 역할을 하고 있다. 본 연구는 광고 서사 수준이 소셜 미디어 플랫폼에서의 전파 효과에 어떻게 영향을 미치는지 탐구하는 것을 목적으로 한다. 광고 데이터에 대한 대규모 분석을 통해 본 연구는 서사 수준과 광고의 바이럴성 지표(전파 횟수, 조회수, 전파 속도 등) 간의 관계를 검토하였다. 관계 연구 방법은 정량적 서사 특성의 내용 분석을 결합하여 서사 수준이 바이럴 전파에 미치는 예측 결과를 평가하였다.

예상 연구 결과에 따르면 질 높은 서사는 광고의 바이럴 효과와 유의미한 상관관계가 있으며, 다양한 서사 유형(감정 기반 및 이야기 기반)이 전파 효과에 미치는 영향이 다르게 나타났다. 이러한 발견은 광고학 및 커뮤니케이션 이론을 풍부하게 할 뿐만 아니라 효과적인 소셜 미디어 광고 전략 수립에도 중요한 의미를 제공한다.

파트너십 구축: 공동 브랜딩에서 성공을 이끄는 요인

진연여, 최재원
순천향대학교 경영학과, 순천향대학교 경영학과
e-mail : chenyanru@sch.ac.kr, jaewonchoi@sch.ac.kr

Forging Partnerships: Unveiling the Success Drivers in Co-Branding

Yanru Chen, Jaewon Choi
Soonchunhyang University, Department of Business Administration,
Soonchunhyang University, Department of Business Administration

요 약

A brand is the most valuable asset of a company that seeks commercial success. As the market becomes increasingly competitive and consumers grow more demanding, business managers are constantly looking for ways to enhance the customer experience and drive business growth. This has led many managers to choose co-branding strategies. However, not all co-branding efforts are successful.

This paper delves into the key factors that contribute to the success of co-branding initiatives, specifically by examining the various attributes and characteristics of co-branded products. It aims to provide a comprehensive analysis of how different product features influence the effectiveness and market performance of co-branding partnerships. Through this exploration, the study seeks to identify the primary drivers that determine whether co-branded products resonate positively with consumers and achieve their intended goals.

사용자를 위한 ASMR 마케팅 콘텐츠의 에코챔버 효과: 텍스트 마이닝 관점에서의 분석

르키아, 최재원
순천향대학교 경영학과
e-mail : rokalahlimi93@sch.ac.kr, jaewonchoi@sch.ac.kr

The Echo Chamber Effect in ASMR Marketing Content for Users: AN Analysis from a text Mining Perspective

Rkia lahlimi, Jaewon Choi
Soon chun hyang University , Department of Business Administration

요 약

Recently, social media content is growing strongly, among them Autonomous sensory meridian response (ASMR) became one of the most trending contents on YouTube. ASMR is widely applied in the field of marketing such as commercial advertising or influencer content. This study aims to explore the echo chamber effect of ASMR marketing content on online users. First, we applied sentiment analysis and topic modeling (LDA) to uncover the user's attitude toward three types of ASMR content such as mukbang, tourism and sponsor content. Then we used social network analysis and multilayer perceptron to explore the difference in communication effect of each ASMR content field. Through the result sentiment analysis and topic modeling, we found online users have better attitudes and responses towards ASMR tourism content than ASMR mukbang and ASMR sponsor content. The results of social network analysis and multilayer perceptron show the ASMR Mukbang has the best communication effect among them, next is ASMR sponsor content, and ASMR tourism content has a weaker communication effect than others. This study provides information to establish an efficient advertising strategy by using ASMR content as a marketing strategy.

도덕 결여 행동과 브랜드 위기: 소셜 미디어 인플루언서 스캔들이 소비자 신뢰에 미치는 영향

이맹룡, 최재원
순천향대학교 경영학과, 순천향대학교 경영학과
e-mail : limengling@sch.ac.kr, jaewonchoi@sch.ac.kr

Moral failure and brand crisis: the impact of social media influencer scandals on consumer trust

Mengling Li, Jaewon Choi
Soonchunhyang University, Department of Business Administration,
Soonchunhyang University, Department of Business Administration

요 약

시대의 발전과 함께 디지털화가 본격화되고 소셜 미디어는 사람들의 삶에 없어서는 안 될 부분이 되었다. 그것은 우리가 다른 사람들과 소통하는 방식을 바꿀 뿐만 아니라 우리의 라이프스타일과 사회적 가치에 큰 영향을 미치며 그 영향력은 점차 다양한 플랫폼에서 확대되고 있다. 소셜 미디어 인플루언서 마케팅도 점차 다양한 브랜드의 마케팅 계획에 포함되었다.

소셜 미디어 인플루언서는 소비자와 소통하는 도구로 여겨진다. 브랜드는 소비자 사이에서 소셜 미디어 인플루언서의 막대한 영향력을 중시하며 인플루언서의 지지가 브랜드 제품 판매를 촉진하고 브랜드의 시장 영향력을 강화할 수 있기를 바란다. 소셜 미디어 인플루언서도 양질의 브랜드와 협력하여 가치를 높이려고 한다. 그러나 브랜드는 혜택을 얻지만 위험도 있다. 브랜드를 지지할 아티스트나 운동선수를 선택하는 것이 도덕적 문제, 불법 문제, 전문적 역량 문제와 같은 위험에 직면하는 것처럼 소셜 미디어 인플루언서도 사회 질서와 좋은 관습을 위반하는 일을 하면 대중의 저항에 부딪혀 브랜드 지지로 이어질 것이다. 브랜드 역시 여론의 폭풍에 휩싸였다. 지속적인 비윤리적 행동에 대한 여론의 영향 외에도, 미디어에 다시 한번 드러난 과거의 비윤리적 역사가 추천되는 제품에도 영향을 미칠 지도 모른다. 반면 소셜 미디어는 거짓말과 중상모략이 쉽게 퍼질 수 있는 환경이다. 때로는 소셜 미디어에서 옳고 그름을 구별하기 어려울 때가 있다. 소셜 미디어에서 대중은 쉽게 직접 조종될 수 있으며, 소비자는 다양한 방식으로 오도되어 사건에 대한 잘못된 인식을 낳을 수도 있다.

본 연구는 소셜 미디어 인플루언서들 사이의 비윤리적 행동과 소비자가 선호하는 브랜드가 소비자의 판단에 어떤 영향을 미치는지 살펴본다. 변수는 소셜 미디어의 확산 속도와 온라인 여론의 영향이다.

서버리스 컴퓨팅에서 콜드 스타트 문제를 다루기 위한 기법들에 대한 분석

조민규, 박상오

중앙대학교 컴퓨터공학과

e-mail : mgjo@cslab.cau.ac.kr, sopark@cau.ac.kr

Analysis of Techniques for Addressing the Cold Start Problem in Serverless Computing

Mingyu Jo, Sangoh Park

Chung-Ang University, School of Computer Science and Engineering

요 약

점점 더 많은 서비스가 자체 인프라에서 Amazon Web Services, Google Cloud Platform, Microsoft Azure와 같은 클라우드 서비스로 이전됨에 따라 클라우드 사용자와 클라우드 공급자가 클라우드를 효율적으로 사용하고 관리할 수 있는 방법에 대한 연구가 활발하게 진행되고 있다. 서버리스 패러다임은 이 새로운 방향에서 새롭게 등장하는 개념이다. 서버리스는 컨테이너를 적극적으로 활용하여 사용자가 원하는 기능에 초점을 맞추기 때문에 클라우드 사용자는 "진정으로" 사용하는 것에 대해서만 비용을 지불하고 클라우드 공급자는 리소스 관리에 집중할 수 있다. 그러나 서버리스 컴퓨팅은 컨테이너 기반이기 때문에 컨테이너의 초기 시작 지연으로 인한 처리 시간이 증가하는 문제가 있다. 이러한 문제를 콜드 스타트 문제라고 하며 본 논문에서는 콜드 스타트 문제를 해결한 다양한 기술을 분석한다.

1. 서 론

서버리스 컴퓨팅은 클라우드 분산 컴퓨팅의 새로운 패러다임으로 부상하고 있다. 서버리스 컴퓨팅 시장 규모는 2025년까지 250억 달러로 성장할 것으로 예상되며, 전 세계 기업의 50%가 서버리스 컴퓨팅을 도입할 것으로 전망된다[1]. 이러한 시장 활성화로 인해 연구 분야에서도 활발히 연구가 진행되고 있다.

서버리스 컴퓨팅은 2014년 Amazon Web Services가 Lambda 서비스를 출시하면서 주목받기 시작했으며, 현재는 Google Cloud Function, Microsoft Azure Functions, Cloudflare Worker 등 다양한 클라우드 제공업체에서 서비스를 제공하고 있다.

커뮤니티에서는 OpenFaaS, OpenWhisk, Kubless, Knetic, OpenLambda 등 Kubernetes 기반 플랫폼이 오픈 소스로 활발히 연구 및 개발되고 있다.

기존 클라우드의 서버풀 방식의 경우 부하를 예측하여 인스턴스를 미리 띄워놓거나 처리할 수 없는 부하가 생길 경우 새로운 인스턴스를 생성하는 방식으로 요청을 처리하여 리소스가 낭비되거나 반응성이 낮아지는 문제가 있다.

서버리스는 클라우드 사용자의 각 요청에 마다 독립적인 새로운 인스턴스를 생성하고 가상 머신 및 컨테이너와 같은 가상화 기술을 통해 요청을 처리한다. 이러한 특성으로 인해 클라우드 사용자는 요청 처리 로직에 집중할 수 있으며, 클라우드 제공업체는 사용자의 분산 요청을 처리할 수 있는 시스템 및 인프라에 집중할 수 있다.

그러나 가상화로 인해 관측의 어려움, 휘발성 등 다양한

문제가 발생하고 있고, 특히 가장 큰 문제는 각 요청에 대해 새로운 가상 인스턴스를 생성하는 데 소요되는 시간으로 인한 지연이다. 이 문제를 콜드 스타트 문제라고 하며, 이를 해결하기 위한 다양한 솔루션이 제안되고 있다[2]. 본 논문에서는 이러한 솔루션들을 분석한다.

본 논문의 2장에서는 서버리스 컴퓨팅에서의 다양한 콜드 스타트 해결 방법을 분석하고, 3장에서는 결론을 제시하고 향후 연구 방향을 제안한다.

2. 관련 연구

콜드 스타트 문제를 해결하기 위해 다양한 기법들이 연구되었다. 크게 캐시를 활용하는 기법과 캐시를 활용하지 않는 기법으로 나눌 수 있다.

빈번하게 사용되는 부분을 캐싱하여 관리하는 연구에는 컨테이너 콜드 스타트 시 장시간이 소요되는 네트워크 초기화 부분을 사전에 로드하여 지연 시간을 감소시키는 기법[3], 자주 사용되는 이미지의 리소스를 해제하지 않고 지속적으로 사용하는 기법[4], 그리고 라이브러리 단위와 같이 빈번하게 사용되는 부분만을 선별적으로 캐싱하는 기법[5] 등이 연구되었다. 이러한 기법들은 캐싱을 위해 많은 리소스를 사용해야 한다는 단점이 있다.

캐시 기법에서 많은 리소스를 사용하는 단점을 극복하기 위해, 가상화 수준을 다양화하는 기법이 제안되었다[6]. 또한, 순간적으로 증가하는 부하를 처리하기 위해 요청을 클러스터링하여 하나의 인스턴스에서 일괄 처리하는 기법이 제안되었다[7]. 그러나 이러한 기법들도 중복된 리소스 사용이나 정확한 통신의 어려움으로 인한 오버헤드 증가

등의 단점이 있을 수 있다.

3. 결론

서버리스 컴퓨팅의 콜드 스타트 문제를 해결하기 위해 인스턴스 재사용, 캐시, 가상화 수준 조정 등 다양한 기법들이 제안되었다.

제안된 기법들은 콜드 스타트로 인한 지연 시간을 일정 부분 완화시키는 데 기여하였다. 그러나 각 노드 간 정보를 효율적으로 공유하기 어렵다는 한계로 인해 순간적인 변화에 적응하는 데 어려움이 있다. 이는 서버리스 환경의 동적 특성을 충분히 활용하지 못하는 결과를 초래한다.

따라서, 향후 연구 방향은 이러한 동적 특성을 기반으로 하여 다음과 같은 측면에 초점을 맞출 필요가 있다:

1. 효율적인 통신 방법: 노드 간 정보 공유를 최적화하여 시스템의 전반적인 응답성을 향상시키는 방안을 모색해야 한다.
2. 분산 처리를 통한 콜드 스타트 지연 개선: 부하를 효과적으로 분산시키고 병렬 처리를 통해 전체적인 시스템 성능을 향상시키는 방법을 연구해야 한다.

이러한 접근 방식은 서버리스 컴퓨팅의 확장성과 유연성을 극대화하면서도 콜드 스타트로 인한 성능 저하를 최소화할 수 있는 균형 잡힌 해결책을 제시할 수 있을 것으로 기대된다.

Acknowledgments

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0020632, 2024년 산업혁신인재성장지원사업)

참고문헌

[1] Costello K. The CIO's Guide to Serverless Computing. City, 2020.

[2] Li Y., Lin Y., Wang Y., Ye K. and Xu C. Serverless computing: state-of-the-art, challenges and opportunities. IEEE Transactions on Services Computing, 16, 2 (2022), 1522-1539.

[3] Mohan A., Sane H., Doshi K., Edupuganti S., Nayak N. and Sukhomlinov V. Agile cold starts for scalable serverless. City, 2019.

[4] Wang K.-T. A., Ho R. and Wu P. Replayable execution optimized for page sharing for a managed runtime environment. City, 2019.

[5] Liu X. Z., Wen J. F., Chen Z. P., Li D., Chen J. K., Liu Y., Wang H. Y. and Jin X. FasSLight: General Application-level Cold-start Latency Optimization for Function-as-a-Service in Serverless Computing. Acm Transactions on Software Engineering and

Methodology, 32, 5 (Jul 2023), 1-29.

[6] Akkus I. E., Chen R., Rimac I., Stein M., Satzke K., Beck A., Aditya P. and Hilt V. {SAND}: Towards {High-Performance} serverless computing. City, 2018.

[7] Wu Z., Deng Y., Zhou Y., Li J. and Pang S. FaaSBatch: Enhancing the efficiency of serverless computing by batching and expanding functions. IEEE, City, 2023.

무인 항공기의 협력적 통신에 관한 연구

김동현, 박상오*

중앙대학교 대학원 컴퓨터공학과, 중앙대학교 소프트웨어학부
e-mail : dhkim@cslab.cau.ac.kr, sopark@cau.ac.kr

A Research of Cooperative Communication in UAV

Donghyeon Kim, Sangoh Park*

School of Computer Science and Engineering, Chung-Ang University

요 약

이 연구에서는 무선 환경의 전력 효율성과 신호 품질 문제를 해결하기 위해 무인 항공기(UAV) 네트워크의 협력 통신 기술을 살펴본다. 이 연구는 릴레이 통신과 분산형 다중 노드 전송 전략에 중점을 둔다. 중계 통신은 중계 UAV 노드를 사용하여 네트워크 커버리지를 확장하고 통신 품질을 향상시킨다. 분산 다중 노드 전송은 공간적 다양성을 활용하여 안정성과 에너지 효율성을 개선한다. 이러한 기술은 전력 소비와 신호 저하 문제를 효과적으로 완화한다. 이번 연구 결과는 협력적 통신 전략이 UAV 네트워크 성능을 향상시켜 보다 강력하고 다양한 항공 통신 시스템을 구현하는 데 매우 중요하다는 것을 보여준다. 이 연구는 재난 대응, 환경 모니터링, 정밀 농업 등 다양한 분야의 UAV 애플리케이션 발전에 기여할 것으로 기대된다.

1. 서 론

최근 몇 년간 배터리 및 모터 기술의 급속한 발전으로 무인 항공기(UAV)의 활용이 크게 확산되고 있다. 이러한 기술적 진보는 데이터 수집 방식을 혁신적으로 변화시켰으며, 정적 센서 배치에서 동적 다차원 수집에 이르기까지 다양한 응용 분야에서 UAV의 활용도가 높아지고 있다. UAV는 농업, 재난 관리, 환경 모니터링, 도시 계획 등 다양한 분야에서 중요한 도구로 자리잡고 있으며, 이에 따라 UAV 네트워크의 효율적인 운용이 더욱 중요해지고 있다.

그러나 UAV의 특성상 배터리에 의존하는 운용 방식으로 인해, 무선 통신에는 여전히 중요한 과제들이 존재한다. 특히 전력 효율성과 최적의 대역폭 활용이 주요 관심사로 떠오르고 있다. UAV의 제한된 배터리 용량은 비행 시간과 데이터 전송 능력에 직접적인 영향을 미치므로, 에너지 효율적인 통신 방식의 개발이 필수적이다.

무선 환경에서의 통신은 여러 가지 도전 과제를 안고 있다. 장애물과 간섭으로 인해 신호 품질이 심각하게 저하될 수 있어, 정확한 데이터 전송이 어려워질 수 있다. 또한 대기 중의 다양한 기상 조건, 전자기적 간섭, 다중 경로 페이딩 등의 요인들이 통신 품질에 부정적인 영향을 미칠 수 있다. 이러한 환경적 요인들은 UAV 네트워크의 안정성과 신뢰성을 저해하는 주요 원인이 된다. 더불어, 데이터 처리와 정밀한 통신을 위한 UAV의 연산 요구는 상당한 에너지 소비를 초래한다. 고해상도 이미지나 비디오 스트리밍, 실시간 센서 데이터 처리 등 복잡한 작업을 수행하기 위해서는 강력한 연산 능력이 필요하며, 이는

곧 높은 에너지 소비로 이어진다. 따라서 연산 효율성과 에너지 소비 사이의 균형을 맞추는 것이 UAV 네트워크 설계의 중요한 과제 중 하나이다.

이러한 다양한 문제들을 해결하기 위해, 여러 노드 간의 협업을 통해 통신 효율성을 높이고 에너지 사용을 최적화하는 협력 통신이 유망한 해결책으로 떠오르고 있다 [1]. 협력 통신 기술은 네트워크 내의 여러 노드들이 서로 협력하여 통신 성능을 향상시키는 방식으로, 단일 노드의 한계를 극복하고 전체 네트워크의 효율성을 높일 수 있는 잠재력을 가지고 있다.

본 연구에서는 UAV 네트워크에서의 협력 통신 기술, 특히 중계 통신과 분산 다중 노드 전송 기술에 초점을 맞추어 살펴볼 것이다. 이를 통해 UAV 네트워크의 성능 향상 가능성을 탐구하고, 향후 연구 방향에 대한 통찰을 제공하고자 한다.

2. 무인 항공기 네트워크에서의 협력적 통신 기술

2.1 중계 통신

무인 항공기(UAV) 네트워크의 릴레이 통신은 중계 노드를 통한 데이터 전송을 통해 네트워크 성능을 크게 향상시키는 중요한 협력 통신 패러다임이다.[2] 이 접근 방식은 네트워크의 커버리지 반경을 확장하고 강력하고 효율적인 데이터 전송 메커니즘을 보장하여 전반적인 통신 품질을 향상시킨다.

이 프레임워크에서 UAV는 중간 노드로서 기능하며, 프로토콜에는 증폭 후 전달(AF) 방식과 디코드 후 전달(DF) 방식의 두 가지 유형이 있다. AF에서는 릴레이 노드가 수

신된 신호를 증폭한 후 재전송한다. 이 방식은 단순성과 속도를 제공하지만 노이즈 증폭의 위험이 있다. DF는 신호를 디코딩, 재인코딩 및 전송하는 과정을 포함한다. 잠재적으로 신호 무결성을 향상시킬 수 있지만 지연 시간이 추가로 발생할 수 있다.

UAV 기반 통신 시스템에 릴레이 노드를 전략적으로 배치하면 이전에는 접근이 불가능하거나 까다로운 환경에서도 연결성을 확보할 수 있다. 또한 이 접근 방식은 거리에 따라 전송 전력 소비가 기하급수적으로 증가하는 무선 네트워크에서 효과적인 통신 거리를 줄여 에너지 효율을 향상시킨다. 이러한 에너지 최적화는 임무 기간과 전체 시스템 성능에 큰 영향을 미치는 UAV 운용에 매우 중요하다.

2.2 분산 다중 노드 전송

다중 입력 단일 출력(MISO) 통신이라고도 하는 분산 다중 노드 전송은 여러 전송 노드를 활용하여 전체 시스템 성능을 향상시키는 고급 형태의 협동 통신 기술이다.[3] 이 패러다임에서는 두 대 이상의 UAV가 동시에 하나의 수신 노드로 신호를 전송하고, 수신된 데이터를 활용하여 정확한 정보를 도출한다.

이 접근 방식을 뒷받침하는 핵심 개념은 공간적 다양성으로, 다중 송신 UAV를 통해 수신 신호의 전파 경로를 다양화함으로써 달성할 수 있다. 양방향 전송 경로를 통해 도착하는 신호는 다양한 정도의 페이딩 및 색도잉 효과를 받게 된다. 이러한 다양성을 활용하면 보다 정확한 데이터 수신이 가능하여 통신 안정성을 크게 향상시킬 수 있다. 결과적으로 이 접근 방식은 통신 링크의 전반적인 견고성을 강화하여 까다로운 환경 조건에서도 신호 수신에 성공할 확률을 높인다.

이 접근 방식은 정확한 데이터 수집과 전송 전력 감소를 통해 에너지 효율을 향상시킨다. 공간적 다양성을 통해 전송 노드 간에 에너지 자원을 효율적으로 분배할 수 있어 시스템의 전체 에너지 소비를 줄일 수 있다.

3. 결론

본 연구에서는 무인 항공기(UAV) 네트워크의 성능을 향상시키기 위한 협력 통신 기술에 대해 중점적으로 살펴 보았다. 특히 릴레이 통신과 분산형 다중 노드 전송 기술에 초점을 맞추어 분석하였다. 이러한 전략들은 네트워크 커버리지를 확장하고 통신 품질을 개선하며 UAV 운영에서 에너지 효율성을 최적화하는 데 상당한 잠재력을 보여 주고 있다.

릴레이 통신 기술, 분산형 다중 노드 전송 기술과 같은 협력 통신 기술들은 UAV 네트워크가 직면한 주요 과제인 전력 소비와 신호 저하 문제를 효과적으로 해결할 수 있는 방안을 제시한다. 이를 통해 더욱 강력하고 다양한 항공 통신 시스템을 구현할 수 있게 되며, 이는 UAV의 활용 범위를 크게 확대시킬 수 있는 잠재력을 가지고 있다.

그러나 이러한 기술들의 실제 구현에는 여전히 몇 가지 과제가 남아있다. 다양한 환경 조건에서의 최적 성능을 위한 조정 알고리즘의 개발, 그리고 다양한 협력 통신 기술들을 통합한 하이브리드 시스템의 설계 등이 필요하다. 따라서 향후 연구에서는 이러한 과제들을 해결하기 위한 노력이 필요하다. 머신러닝과 인공지능 기술을 활용한 적응형 알고리즘의 개발, 다양한 협력 통신 기술들을 효과적으로 결합한 하이브리드 시스템의 연구, 그리고 실제 운용 환경에서의 성능 검증 등이 중요한 연구 주제가 될 것이다. 이를 통해 다양한 운용 시나리오에서 UAV 네트워크의 성능을 더욱 최적화하고, 궁극적으로는 UAV 기술의 실용성과 효용성을 크게 향상시킬 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

이 연구는 2024년 정부(산업통상자원부)의 재원으로 한국 산업기술진흥원의 지원을 받아 수행된 연구임. (P0020632, 2024년 산업혁신인재성장지원사업).

참고 문헌

- [1] W. Su, A. K. Sadek and K. J. Ray Liu, "Cooperative Communication Protocols in Wireless Networks: Performance Analysis and Optimum Power Allocation", vol. 44, no. 2, pp. 181 - 217, Jan. 2008, doi: 10.1007/s11277-007-9359-z.
- [2] L. J. Natonski, T.-H. Kim and B. Hadden, "The Effect of Relay Node and Power Control on Performance in Multi-hop Wireless Network", pp. 540 - 544, Jun. 2017, doi: 10.1109/MASS.2017.99.
- [3] M. A. T. Alrubei, I. A. Alshimaysawe, A. N. Hassan and A. H. K. Khwayyir, "Capacity analysis & performance comparison of SISO, SIMO, MISO & MIMO systems", vol. 1530, no. 1, p. 012077, May 2020, doi: 10.1088/1742-6596/1530/1/012077.
- [4] M. Champion, P. Ranganathan and S. Faruque, "UAV swarm communication and control architectures: a review", vol. 7, no. 2, pp. 93 - 106, Jun. 2019, doi: https://doi.org/10.1139/juvs-2018-0009.
- [5] Y. Zhou, B. Rao and W. Wang, "UAV Swarm Intelligence: Recent Advances and Future Trends," in IEEE Access, vol. 8, pp. 183856-183878, 2020, doi: 10.1109/ACCESS.2020.3028865.
- [6] M. F. J. Pinkney, D. Hampel, and S. DiPierro, 'Unmanned aerial vehicle (UAV) communications relay', vol. 1. IEEE, pp. 47 - 51, 1996.

디지털 트윈을 활용한 고부가가치 저차원 소재의 물성 예측

이재원, 양승화
중앙대학교 에너지시스템공학과
e-mail : herculepoirot163@naver.com

Property Prediction of High-Value Low-Dimensional Materials Using Digital Twin Technology

Jaewon Lee, Seunghwa Yang
School of energy system engineering, Chung-Ang University

요 약

디지털 트윈 기술의 발전은 기존의 시스템 건전성 진단을 위한 가상모델을 넘어 소재 개발 및 공정을 아우르는 전주기적인 디지털 가상화로 나아가고 있다. 특히 저차원 나노소재 기반의 복합재료를 설계하는데 있어 인실리코 해석 기반의 소재-물성 데이터베이스 구축은 필수불가결한 과정이 되었으며, 해당 데이터를 기반으로 머신러닝과의 연계를 통한 프레임워크 역시 제시되고 있다. 이를 위해 다양한 소재에 대한 구조-물성-환경 상관관계가 연구되고 있으나, 아직까지 통합적인 가상화를 위한 데이터베이스는 부족한 실정이다.

본 연구에서는 분자동역학 전산모사를 통해 다기능성 고부가가치 나노소재인 질화붕소 나노튜브의 반경 변화에 따른 압전탄성 물성의 변화를 규명하였다. 붕소와 질소의 원자간 포텐셜 에너지 모사를 위해 강체이온근사와 Tersoff 포텐셜을 채택하였으며, 선형 압전탄성 구성방정식을 기반으로 일축인장에 따른 전기적 변위와 응력을 산출하여 압전상수와 응력을 예측하였다. 그 결과 나노튜브의 대표적인 구조적 특징인 키랄리티가 압전탄성 물성에 유의미한 영향을 미치는 것을 확인하였으며, 이를 해석 결과에 기반한 경험적 수식 모델로 제시하였다. 또한 이러한 구조-물성 상관관계를 나노튜브와 나노시트의 결합에너지와 결합길이 변화의 차이로 설명하였다.

본 논문은 분자동역학 전산모사를 통해 질화붕소 나노튜브의 구조-물성 상관관계를 경험적 수식으로 제시함으로써 신뢰도 높은 소재 설계와 가상화를 위한 통찰을 제공할 것으로 예측된다.

인실리코 해석을 이용한 방열복합재료의 물성 예측

기나현, 양승화
중앙대학교 에너지시스템공학과
e-mail : yr1375@naver.com

A Study of Heat Dissipation Composite Using In-Silico Analysis

Nahyun Ki, Seunghwa Yang
School of Energy system Engineering, Chung-Ang University

요 약

본 연구는 ICT(정보통신기술) 분야에서 발생하는 열 관리 문제 해결을 위한 방열 복합 재료 물성 예측 방법론을 제안하였다. 고방열 소재는 IoT, 인공지능을 활용하기 위한 고성능 반도체, ICT 기기 등 그 발전 방향과 적용 분야가 매우 다양하다. 특히, 기술의 발전으로 인해 진행된 부품의 경량화·소형화는 고도화된 방열 시스템 구축과 함께 보다 향상된 물성의 재료를 필요로 한다. 이러한 고성능 물성 예측을 위해 본 연구는 인실리코 해석을 활용한 방열 복합재료의 물성 예측을 주제로, 고전적인 물성 예측 방법인 실험적 접근 방법에서 벗어난 컴퓨터 시뮬레이션을 수행하였다.

가교 구조를 가진 에폭시 고분자와 열전도성 필러가 포함된 복합재료를 분자동역학 전산모사 기반의 Material Studio 프로그램을 통해 모델링하였다. 이후, 클러스터를 기반한 가상 컴퓨팅 환경에서의 등온-등압 (NPT) 앙상블을 통해 상온 대기압 조건에서의 안정화를 진행하고, 이후 다양한 시뮬레이션 계산을 통해 여러 변수와 환경에 따른 기계적 물성 및 열전도도를 예측하였다. 이러한 시뮬레이션 방법론은 온도, 압력과 같은 변수 제어가 실험적 접근보다 엄격한 통제가 가능하며, 실험을 통한 물성 예측이 가진 시간과 비용의 한계를 극복할 수 있다.

해당 연구는 인실리코 기반 시뮬레이션 방법론 적용을 통해 재료의 설계 단계부터 최적화된 물성 예측의 가능성을 제시한다. ICT 기기의 소형화와 고성능화로 인한 열 문제 해결 및 고기능성의 방열 재료 설계에 기여하며 정보통신기술 산업에서 필수적인 방열 솔루션을 제시에 필요한 기초 자료로써 활용될 수 있을 것이다.

2024년 소비 트렌드를 반영한 포스트 휴먼 시대에 AI를 활용한 과학기술 정책전략 수립 및 디지털 대전환 방안에 관한 연구

최장원, 황명권, 최선희
한국과학기술정보연구원 정책전략본부 정책연구센터
e-mail : jwchoi@kisti.re.kr, mgh@kisti.re.kr, sunny.choi@kisti.re.kr

Research on the establishment of science and technology policy strategy and digital transformation plan using AI in the post-human era reflecting consumption trends in 2024

Choi Jang Won, Myunggwon Hwang, Choi Seon Heui
Korea Institute of Science & Technology Information, Div. of Policy and Strategy, Policy Research Center

요약

본 논문은 2024년 10대 소비 트렌드 분석을 통한 미래 사회 변화에 대비하여 포스트 휴먼 시대의 디지털 대전환을 위한 전략적 시사점과 장기적 관점에서의 과학기술 및 AI 정책 방향 설정을 위한 정책 제언과 향후 연구과제를 제시한다.

1. 서론: 포스트 휴먼 시대의 도래와 디지털 콘텐츠의 역할 변화

2024년 현재, 인공지능(AI) 기술의 발전은 인간의 능력을 뛰어넘는 수준에 도달하며 포스트 휴먼 시대를 열고 있다. 이러한 변화는 과학기술 및 디지털 콘텐츠의 역할에도 큰 영향을 미치고 있다. 과거에는 정보 전달 도구였던 디지털 콘텐츠는 이제 인간과 상호작용하며 새로운 경험을 제공하는 역할을 수행하게 되었다[1]. AI 기술의 발전과 디지털 콘텐츠의 역할 변화는 사회, 경제, 문화 전반에 걸쳐 긍정적, 부정적 영향을 미칠 수 있다. 따라서 이러한 변화에 효과적으로 대응하기 위한 과학기술 정책전략 수립 및 디지털 대전환 방안 마련이 시급하다. 본 연구는 AI 기술을 활용한 과학기술 정책전략 수립 및 디지털 대전환 방안에 대한 연구를 통해 포스트 휴먼 시대의 사회적, 경제적 가치 창출 방안을 모색하고자 한다.

2. 관련 연구

2.1 포스트 휴먼 시대의 과학기술 및 AI 현황 분석

2024년 현재 AI 기술은 급격한 발전을 이루고 있으며, 머신러닝, 딥러닝 기술의 발전은 자동화, 예측, 의사결정 등 다양한 분야에서 활용되고 있다. 특히, 자율주행, 로봇공학, 생명공학 분야는 AI 기술과의 결합을 통해 혁신적인 변화를 맞이하고 있다. AI 기술의 발전은 인간과의 공존, 윤리적 문제, 법적 규제 등 다양한 이슈를 제기한다. 특히, AI 기술의 윤리적 사용과 규제는 포스트 휴먼 시대의 주요 과제로 떠오르고 있다[2].

2.2 디지털 콘텐츠의 대전환과 AI 기술의 역할

AI 기술은 디지털 콘텐츠 산업 전반에 걸쳐 적용되고 있으며, 새로운 콘텐츠 형식과 사용자 경험을 창출하는 데 중요한 역할을 하고 있다. 예를 들어, 맞춤형 콘텐츠 추천, 인터랙티브 콘텐츠, 가상현실(VR) 및 증강현실(AR) 등은 AI 기술을 기반으로 발전하고 있다[3]. AI 기술은 사용자 개개인에게 맞춤형 콘텐츠를 제공하고, 사용자와의 상호작용을 통해 새로운 경험을 제공한다.

2.3 AI를 활용한 과학기술 정책전략 수립의 중요성

AI 기술 발전은 사회 전반에 걸쳐 긍정적, 부정적 영향을 미칠 수 있다. 긍정적인 측면으로는 생산성 향상, 의료 서비스 개선, 교육 기회 확대 등이 있고, 부정적인 측면으로는 일자리 감소, 사회적 불평등 심화, AI 기술 악용 등이 있다[4]. AI 기술의 윤리적 사용, 프라이버시 보호, 데이터 보안 문제는 포스트 휴먼 시대의 중요한 과제이다. AI 기술의 윤리적 사용을 위한 국제적 협력과 법적 규제가 필요하며, 개인 정보 보호 및 데이터 보안 강화 방안 마련이 시급하다[2].

3. 정책 제언

3.1 포스트 휴먼 시대의 디지털 대전환을 위한 전략적 시사점

AI 기술은 포스트 휴먼 시대에 사회, 경제, 문화 전반에 걸쳐 긍정적, 부정적 영향을 미칠 수 있다. 따라서 AI 기술의 윤리적 사용, 프라이버시 보호, 데이터 보안 문제를 해결하고, AI 인재 양성, 디지털 인프라 구축 등을 위한 정책적 지원 등이 필요하다. 구체적 예시로 AI 기술을 활

용한 과학기술 정책 수립 및 디지털 대전환을 통해 사회적, 경제적 가치 창출, AI 기술의 윤리적 사용, 프라이버시 보호, 데이터 보안 문제 해결, AI 인재 양성 및 디지털 인프라 구축을 위한 정책적 지원, 인간과 AI의 공존을 위한 사회적 합의 및 윤리적 가이드라인 마련 등이다.

3.2 장기적 관점에서의 과학기술 및 AI 정책 방향 설정

포스트 휴먼 시대의 과학기술 및 AI 정책은 인간과 AI의 공존을 기반으로 인간 중심의 가치를 실현하는 방향으로 설정되어야 한다. 또한, 장기적 관점에서 지속 가능한 발전을 위한 연구개발 투자 확대와 사회적 합의를 이루는 노력이 필요하다.

(표 1) 디지털 대전환을 위한 정책 제언

정책 제언	정책 예시
정부와 민간 부문의 협력을 통한 AI 기술 R&D 지원 확대	- 핵심기술 개발 및 기초연구 지원 확대 - AI 기술 상용화 및 산업화를 위한 정부와 민간의 협력 체계 구축 - AI 기술 인재 양성 및 교육 프로그램 개발
AI 인재 양성 및 교육 시스템 개선	- AI 관련 교육 과정 확대 및 교육 수준 강화 - AI 기술 전문 인력 양성을 위한 산학 협력 프로그램 개발 - 평생 교육 시스템을 통한 AI 기술 교육 기회 제공
디지털 경제 성장을 위한 인프라 구축과 법적·윤리적 기반 마련	- 초고속 인터넷 및 5G 네트워크 구축 확대 - AI 기술 관련 법적 규제 및 윤리 가이드라인 마련 - 데이터 개방 및 활용 정책 마련

(표 2) 소비 트렌드 분석[5]을 통한 과학기술 및 AI 정책 제언

정책 제언	정책 예시
분초사회 트렌드 반영: 시간 효율성을 높이는 AI 기술개발 및 활용 정책 추진	- AI 기반 시간관리 도구 개발 및 보급 - 맞춤형 정보 제공 및 업무 자동화 시스템 구축 - 시간 활용 교육 및 컨설팅 제공
호모 프롬프트 트렌드 반영: 인간과 AI의 상호작용 능력 향상 정책 추진	- 프롬프트 엔지니어링 교육 및 인재 양성 - AI 인터페이스 개선 및 사용자 친화적 디자인 개발 - 인간-AI 협업을 위한 교육 및 문화 조성
지속 가능한 발전: AI 기술을 활용하여 환경 문제 해결 및 지속 가능한 발전 목표 달성	- 친환경 에너지 개발 및 사용 확대 - 환경 오염 감소 및 자원 보존 기술 개발 - 지속 가능한 소비 및 생산 문화 조성
국가 경쟁력 강화: AI 기술을 기반으로 새로운 산업 생태계 구축 및 국가 경쟁력 강화	- 미래 성장동력 발굴 및 핵심기술 개발 - AI 기반 혁신기업 육성 및 투자 확대 - 글로벌시장 진출 지원 및 규제 완화

4. 결론 및 향후 연구과제

4.1 결론

본 연구는 포스트 휴먼 시대의 과학기술 정책전략 수립 및 디지털 대전환 방안에 대한 기본적인 틀을 제시하고, 2024년 10대 소비 트렌드를 반영하여 미래 사회 변화에

대비한 정책을 제안했다. 포스트 휴먼 시대는 과학기술과 인공지능이 인간의 삶에 깊이 영향을 미치는 시대이다. 따라서, 과학기술 정책전략 수립 및 디지털 대전환을 통해 미래 사회 변화에 대비하고 국민 삶의 질 향상을 도모하는 것이 중요하다.

4.2 향후 연구과제

(표 3) 향후 연구 과제

향후 연구과제	주요 내용
윤리적 문제 및 사회적 책임에 대한 심층적인 논의를 통한 사회적 합의 도출	AI 기술 발전과 활용에 따른 윤리적 문제 및 사회적 책임에 대한 심층적인 논의, 특히, 개인정보 보호, 편향성 문제, 일자리 감소 등 주요 윤리적 이슈에 대한 해결 방안 마련
각 트렌드에 대한 구체적인 정책 제안 및 실행 방안 마련, 타당성 및 실현 가능성 평가	향후 연구 방향으로는 각 트렌드의 특성을 고려하여 특정 분야별 AI 기술 적용 연구, AI 기술의 사회적 영향 평가 연구, AI 윤리 및 사회적 책임에 대한 연구 등이 있으며, 정책의 예상되는 효과와 부작용에 대한 평가, 관련 이해 관계자들의 의견 수렴 등을 통해 실효성 향상
2024년 10대 소비 트렌드의 심층 분석 및 AI 기술과의 연관성 분석	각 트렌드의 발생 배경, 주요 특징, 사회 경제적 영향 등을 구체적으로 파악. 또한, 각 트렌드와 AI 기술의 연관성을 명확하게 제시하고, AI 기술이 각 트렌드의 변화를 어떻게 가속화하거나 변화시킬 수 있는지에 대한 분석 필요
국제 협력 및 글로벌 파트너십 구축 방안 마련	AI 기술의 윤리적 사용, 프라이버시 보호, 데이터 보안 문제는 국제적인 협력을 통해 해결해야 하는 과제. 글로벌 파트너십 구축을 통해 AI 기술의 긍정적 활용을 촉진하고, 국제 사회의 공동 번영에 기여할 수 있도록 해야 함

참고 문헌

[1] Erich Prem, "Principles of digital humanism: A critical post-humanist view", Journal of Responsible Technology, Volume 17, March 2024
 [2] Daniel Schiff, Justin Biddle, Jason Borenstein, Kelly Laas, "What's Next for AI Ethics, Policy, and Governance? A Global Overview", Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, February 2020
 [3] A May, A Sagodi, C Dremel, B van Giffen, "Realizing Digital Innovation from Artificial Intelligence", ICIS, 2020
 [4] G Dagnaw, "The Future of Work in the AI Era: Challenges and Opportunities", digitalcommons.kennesaw.edu, 2020
 [5] 김난도 외, "트렌드 코리아 2024 : 2024 대한민국 소비 트렌드 전망", 2023

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(K24L5MIC1 : 과학기술데이터 정책전략 연구)

기억재생 나노칩 R&D 로드맵 수립 및 추진전략: 정책적 제언

최장원, 황명권, 최선희
한국과학기술정보연구원 정책전략본부 정책연구센터
e-mail : jwchoi@kisti.re.kr, mgh@kisti.re.kr, sunny.choi@kisti.re.kr

Establishing a Roadmap for Memory Regeneration Nanochip R&D: Policy Recommendations

Choi Jang Won, Myunggwon Hwang, Choi Seon Heui
Korea Institute of Science & Technology Information, Div. of Policy and
Strategy, Policy Research Center

요약

본 논문은 기억재생 나노칩 기술개발을 위한 정책적 제언을 제시한다. 먼저, 체계적인 R&D 로드맵 수립과 유연한 운영이 필요하며, 다학제적 협력을 통해 전문가들의 협력을 촉진하고 연구 성과의 공유를 강화해야 한다. 윤리적 및 사회적 문제 대응을 위해 가이드라인을 마련하고 사회적 수용성을 고려한 연구가 필요하다. 또한, 국제적 협력을 강화하고, 정부의 적극적인 지원과 투자를 통해 연구개발 자금 확대, 인력 양성, 세제 혜택 등 다양한 지원 정책이 필요하다.

1. 서론

인간의 뇌는 여전히 미지의 영역으로 남아있으며, 뇌 기능의 이해와 뇌 질환 치료는 인류가 직면한 가장 큰 과제 중 하나이다. 특히, 치매와 같은 퇴행성 뇌 질환은 환자 개인뿐만 아니라 가족과 사회 전체에 막대한 부담을 초래하며, 이에 대한 효과적인 치료법 개발은 시급한 과제이다 [1]. 세계보건기구(WHO)에 따르면, 전 세계적으로 약 5천만 명이 치매를 앓고 있으며, 2030년에는 8천2백만 명, 2050년에는 1억 5천2백만 명으로 증가할 것으로 예상된다 [1]. 치매는 환자의 기억력, 사고력, 판단력 등 인지기능을 저하시켜 일상생활에 어려움을 초래하며, 결국에는 죽음에 이르게 하는 심각한 질환이다. 치매로 인한 사회적 비용 또한 막대하여, 2019년 기준 전 세계적으로 치매 관리 비용은 약 1조 3천억 달러에 달하는 것으로 추정된다[1]. 최근 뇌과학과 나노 기술의 발전은 기억재생 나노칩과 같은 혁신적인 기술 개발 가능성을 제시하고 있으며, 이는 뇌 질환 치료의 새로운 패러다임을 열 것으로 기대된다. 기억재생 나노칩은 뇌에 이식되어 손상된 뇌 기능을 복원하거나 향상시키는 것을 목표로 하는 혁신적인 기술이다. 이 기술은 뇌 질환 치료뿐만 아니라 인간의 인지 능력 향상에 활용될 수 있는 잠재력을 가지고 있어, 미래 사회에 막대한 영향을 미칠 것으로 예상된다.

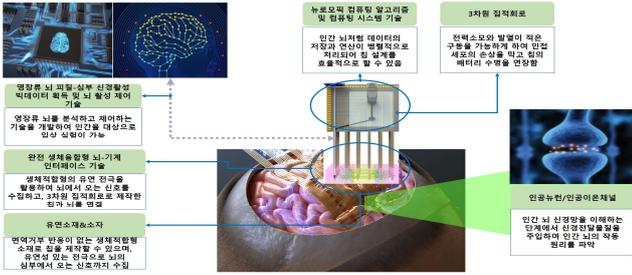
본 논문은 기억재생 나노칩 R&D 로드맵 수립 및 추진 전략의 필요성 및 구체적인 방향성과 정책적 시사점을 제시하고자 한다. 먼저, 기억재생 나노칩 기술의 현황과 전망을 살펴보고, 국내외 R&D 동향을 분석한다. 이를 바탕으로, 성공적인 R&D 로드맵 수립을 위한 핵심 요소들을 도출하고, 정책적 제언을 제시한다.

2. 관련 연구: 기억재생 나노칩 기술의 현황과 전망

현재 기억재생 나노칩 기술은 초기 단계에 있으며, 동물 실험을 통해 가능성을 확인하는 수준이다. 예를 들어, 2013년 미국 서던캘리포니아대학교(USC) 연구팀은 쥐의 뇌에 이식된 칩을 통해 뇌의 특정 부위에 전기 자극을 가하여 기억을 형성하고 저장하는 데 성공하였다[2]. 또한, 2018년 미국 웨이크포레스트대학교 연구팀은 원숭이의 뇌에 이식된 칩을 통해 뇌 활동을 기록하고 분석하여 팔 움직임을 예측하는 데 성공하였다[3]. 뇌과학과 나노 기술의 급속한 발전은 기억재생 나노칩 기술의 실현 가능성을 높이고 있으며, 미래에 상용화되면 교육, 산업 등 다양한 분야에 혁신을 가져올 것으로 기대된다.

미국, 유럽 등 선진국들은 기억재생 나노칩 기술 개발에 적극적으로 투자하고 있다. 특히, 미국의 BRAIN Initiative는 뇌 기능의 이해를 위한 대규모 연구 프로젝트로, 뇌 지도 작성, 뇌 활동 측정 기술 개발 등 다양한 연구를 지원하고 있다[4]. 유럽의 Human Brain Project는 슈퍼컴퓨터를 이용한 인간 뇌 시뮬레이션을 통해 뇌 기능을 이해하고 뇌 질환 치료법을 개발하는 것을 목표로 하고 있다[5].

국내에서도 뇌과학 연구에 대한 투자가 지속적으로 증가하고 있으며, 기억재생 나노칩 기술 개발을 위한 다양한 연구가 진행되고 있다. 한국과학기술연구원(KIST)은 뇌과학연구소를 설립하여 뇌 기능 이해, 뇌 질환 치료, 인공지능 개발 등 다양한 연구를 수행하고 있다. 기초과학연구원(IBS)은 인지및사회성연구단을 선정하여 뇌 기능의 근본 원리를 탐구하고 뇌 질환 치료를 위한 혁신적인 기술 개발을 추진하고 있다.



(그림 1) 기억재생 나노칩 R&D 기획 분야 예시(KISTI)

3. 기억재생 나노칩 R&D 로드맵 수립 및 추진 전략

3.1 핵심 요소 도출

- 명확한 목표 설정: R&D 로드맵은 명확하고 구체적인 목표를 설정해야 한다. 목표는 기술 개발 단계, 성능, 안전성 등 다양한 측면을 고려하여 설정되어야 하며, SMART (Specific, Measurable, Achievable, Relevant, Time-bound) 원칙에 따라 구체적이고 측정 가능하며 달성 가능하고 현실적이며 시의적절해야 한다[6].
- 다학제적 협력: 기억재생 나노칩 기술 개발은 뇌과학, 나노기술, 재료공학, 컴퓨터공학 등 다양한 분야의 전문 지식이 필요하다. 따라서 다양한 분야 전문가들의 협력과 이를 위한 효과적인 협력 체계 구축이 필수적이다[7].
- 윤리적, 사회적 문제에 대한 선제적 대응: 뇌 기능 조작, 개인정보 유출 등 윤리적, 사회적 문제를 야기할 수 있다. 따라서 기술 개발 과정에서 발생할 수 있는 문제점들을 예측하고, 이에 대한 선제적인 대응 방안을 마련해야 한다[8].
- 국제적 협력 연구: 기억재생 나노칩 기술 개발은 장기적이고 대규모 투자가 필요한 분야이다. 정부는 연구개발 자금 지원, 인프라 구축, 인력 양성 등 다양한 지원 정책을 통해 기술 개발을 촉진해야 한다.

3.2 정책적 제언

- R&D 로드맵 수립 및 추진: 기억재생 나노칩 기술 개발을 위한 장기적이고 체계적인 R&D 로드맵을 수립하고, 이를 효과적으로 추진하기 위한 전략을 마련해야 한다. 로드맵은 기술 개발 목표, 단계별 추진 계획, 예산 확보 방안, 성과 평가 지표 등을 포함해야 하며, 정기적인 점검과 수정을 통해 유연하게 운영되어야 한다.
- 다학제적 협력 체계 구축: 다양한 분야 전문가들의 협력을 촉진하기 위한 플랫폼을 구축하고, 연구정보 공유, 공동 연구 수행 등을 지원해야 한다. 또한, 관련 분야 연구자들이 함께 참여하는 융합 연구 프로그램을 활성화하고, 연구 성과의 공유 및 확산을 위한 시스템을 구축해야 한다[18].
- 윤리적, 사회적 문제 대응: 기억재생 나노칩 기술 개발 및 활용 과정에서 발생할 수 있는 윤리적, 사회적 문제에 대한 가이드라인을 마련하고, 사회적 합의를 도출하기 위한 노력이 필요하다. 기술 개발 초기 단계부터 윤리적, 사회적 문제를 고려하여 연구를 진행하고, 시민 참여를 통해 사회적 수용성을 높여야 한다[8, 9].
- 국제적 협력 연구 강화: 선진국들과의 공동 연구, 기술

교류 등을 확대하고, 국제 컨소시엄 참여를 통해 글로벌 협력 네트워크를 구축해야 한다. 또한, 국내 연구자들의 해외 연구 기관 파견, 국제 학술 대회 개최 등을 통해 국제적인 연구 협력을 강화해야 한다.

- 정부 지원 확대: 기억재생 나노칩 기술 개발에 대한 정부의 적극적인 지원과 투자가 필요하며, 연구개발 자금 지원 확대, 세제 혜택 제공, 인력 양성 프로그램 운영 등 다양한 지원 정책을 마련해야 한다. 특히, 기초연구 분야에 대한 장기적인 투자와 함께, 기술 상용화를 위한 산업계와의 협력을 강화해야 한다.

4. 결론

기억재생 나노칩 기술은 뇌 질환 치료와 인간 능력 향상에 혁신적인 변화를 가져올 수 있는 잠재력을 가지고 있다. 하지만 기술 개발 과정에서 발생할 수 있는 윤리적, 사회적 문제에 대한 선제적인 대응과 함께, 다학제적 협력, 국제적 협력 연구, 정부의 적극적인 지원이 뒷받침되어야 한다. 본 논문에서 제시된 정책적 제언들이 기억 재생 나노칩 기술 개발을 위한 효과적인 R&D 로드맵 수립 및 추진 전략 마련에 기여할 수 있기를 기대한다.

참고 문헌

[1] <https://www.who.int/news-room/fact-sheets/detail/dementia>
 [2] Deadwyler, S. A., et al, "A cognitive prosthesis for memory facilitation by closed-loop functional ensemble stimulation of hippocampal neurons in primate brain", *Experimental Neurology*, Vol. 287, Jan 2017
 [3] Bouton, C. E., et al, "Restoring cortical control of functional movement in a human with quadriplegia", *Nature*, 533, Apr 2016
 [4] <https://braininitiative.nih.gov/>
 [5] Amunts, K., et al, "The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain", *Neuron*, 92, Nov 2016
 [6] Doran, G. T, "There's a S.M.A.R.T. Way to Write Management's Goals and Objectives". *Management Review*, 70, 1981
 [7] Committee on Key Challenge Areas for Convergence and Health, "Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond", Washington(DC), National Academies Press, Jun 2014
 [8] Yuste, R., et al, "Four ethical priorities for neurotechnologies and AI", *Nature*, 551, 2017
 [9] Ienca, M., & Andorno, R, "Towards new human rights in the age of neuroscience and neurotechnology", *Life Sciences, Society and Policy*, 13, 2017

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 자체사업으로 수행된 연구입니다.(J24JR075-24: 기억재생나노칩: 두뇌 보조용 전자 장치)

인공지능 윤리 글로벌 공동연구 네트워크 탐색

유화선*, 꺾영*, 정도범**

한국과학기술정보연구원 정책연구센터*, 한국과학기술정보연구원 조사분석평가팀**
e-mail : hsyoun@kisti.re.kr, young.kwak@kisti.re.kr, dbchung@kisti.re.kr

The Network Analysis of International Collaborative Research in the Field of AI Ethics

Hwasun You, Young Kwak, Do-Bum Chung
Korea Institute of Science and Technology Information(KISTI)

요 약

인공지능(AI) 기술의 발전은 보안 위협, 차별, 편향 등 부작용을 동반하고 있음에 따라 세계 각국과 국제기구는 인공지능 윤리 문제에 대응하기 위해 규범 마련 등 정책적 대응을 본격화하고 있다. 이에 따라, 본 연구에서는 인공지능 윤리와 관련된 국제공동연구 현황 분석을 실시하였다. 분석 결과, 공동 연구를 활발히 수행하고 있는 국가는 미국, 영국, 독일 등으로 소수 국가에서 연구가 집중되고 있으며, 미국을 중심으로 국제공동연구 수행이 이루어지고 있는 것으로 확인되었다. 이러한 분석 결과를 기반으로 시사점을 제시하였다.

1. 서 론

인공지능(AI) 기술의 발전과 이로 인한 제품 및 서비스로의 확장은 사회 전반에 큰 변화를 가져왔다. 인공지능 기술 발전은 인간의 삶을 편리하게 하고 인간이 풀지 못하는 난제를 해결하는 이점도 있으나 보안 위협, 차별, 편향성 등 부작용을 동반하고 있다. 특히 인공지능 확산과 함께 인공지능 사고 건수는 2023년 123건으로 2013년 이후 약 20배나 증가하였다[1]. 이에 따라 세계 각국과 국제기구는 인공지능의 신뢰성, 안전성, 책임성 등 인공지능 윤리 문제를 해결하기 위해 규범 마련 등 정책적 대응을 본격화하고 있다. OECD는 AI 원칙 발표('19.5), UNESCO의 AI 윤리 권고('21.11), 미국은 '안전하고 신뢰할 수 있는 AI에 대한 행정명령' 발표('23.10), EU는 '신뢰할 수 있는 AI 윤리 지침('19.4) 및 AI 법안 통과('24.3) 등 인공지능이 초래하는 부작용을 방지하기 위해 노력하고 있다[2]. 특히 AI 정상회담('23.11)에서 각국 정상들이 AI 안전에 관한 공동선언문을 발표하고, 최근 AI 서울 정상회의('24.5)에서 AI 안전성 강화, 포용, 상생 방안 등을 논의하는 등 인공지능 윤리에 대해 국제적으로 함께 논의되고 있는 실정이다. 따라서 본 연구에서는 국가별 인공지능 윤리에 대한 정책적 연구동향과 국가 간 공동연구 활동을 살펴보고자 하였다.

2. 연구 방법

본 연구는 인공지능 윤리와 관련된 국제공동연구 현황을 살펴보기 위해 논문 데이터를 활용하였다. Web of Science에서 2018년부터 2022년까지 발표된 인공지능 윤리 분야 SCI 논문 DB를 활용하였다. 논문은 분석대상 키워드 기반 검색식 작성, 관련 분야 전문가 검토를 통해 검

색식을 작성하고, 논문 검색 DB 구축, 데이터마이닝 기법을 활용한 데이터 추출 및 분석 순으로 분석용 데이터를 구축하였다. 이를 통해 최종적으로 786건의 논문 데이터를 수집하였다.

분석 방법은 KDD/KM(Knowledge Discovery in Database/Knowledge Map) 방법론을 활용하여 국가 간 공동연구 네트워크 분석을 실시하였다. KDD/KM은 빅데이터로부터 마이닝(mining) 기반의 분석을 통해 유용한 지식을 발견하고, 이를 시각화 하는데 활용되는 정량적 분석 방법론이다. 본 연구에서는 동일한 서지정보에 공저자(국가)가 참여하는 공저자(Co-author) 네트워크 분석을 실시하였다. 본 연구의 실증분석을 위해 통계 패키지는 Netminer를 활용하였다.

3. 분석 결과

3.1 연구활동도

본 연구에서는 우선 인공지능 윤리 분야에서 우수한 연구활동도를 보이고 있는 국가를 살펴보았다. 최근 5년간(2018~2022년) 인공지능 윤리 관련 논문을 대상으로 연구활동도가 높은 상위 10개 국가는 표 1과 같다.

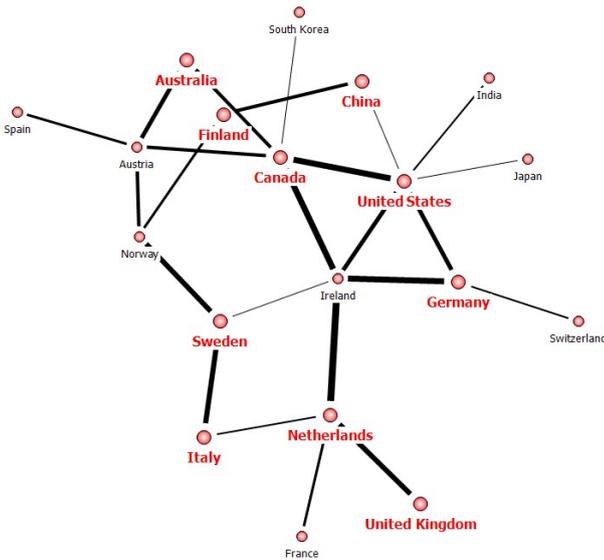
동 기간 상위 10개 국가가 발표한 논문 수는 국가 간 공동연구 논문 수를 고려하여 786건으로 나타났으며, 미국이 187건 논문을 발표하여 상위 120개국 중 총 논문의 23.79% 비중을 차지하였다. 2위는 영국으로 97건의 논문을 발표하여 12.34% 비중을 차지하였으며, 독일, 핀란드, 캐나다, 중국 등을 포함하여 소수 국가에서 연구가 집중되어 있는 경향을 보였다.

(표 1) 연구활동도 글로벌 상위 10개 국가

순위	국가	논문 수	비중(%)
1	United States	187	23.79
2	United Kingdom	97	12.34
3	Germany	67	8.52
4	Finland	50	6.36
5	Canada	49	6.23
6	China	49	6.23
7	Netherlands	45	5.73
8	Australia	35	4.45
9	Italy	24	3.05
10	Sweden	21	2.67

3.2. 공동연구 네트워크 현황

본 연구에서는 KDD/KM 방법론 중 저자 상호인용 방법을 활용하여 국가 간 공저논문 기반 공동연구 네트워크 분석을 실시하였다. 아래 시각화한 그림에서 노드 크기는 국가 간 공동연구 논문 수, 링크 굵기는 공동연구 빈도를 의미한다. 국가 간 공동연구는 미국, 캐나다, 독일, 영국, 네덜란드 순으로 높게 나타났으며, 미국은 캐나다와의 협력에 중점을 두고 활발한 공동연구를 진행하고 있는 것을 확인하였다.



(그림 1) 국가 간 공동연구 네트워크 현황

국가 간 공저논문 네트워크의 국가별 중심성(연결, 매개, 근접) 분석을 실시하였으며, 그 결과는 표 2와 같다. 분석 결과, 국가 간 국제공동연구 중심성 값이 모두 최상위 국가는 미국으로 나타나 네트워크 내에서 영향력이 높다고 할 수 있다. 연결중심성은 특정 개체가 사회 시스템 내부의 다른 개체들과 직접적으로 맺은 연결이나 관계의 수를 측정하며 공저논문 수와 비율을 파악하는데 유용하므로[3], 아래 표는 연결중심성을 기준으로 상위 국가를 작성하였다. 연결중심성은 미국, 캐나다, 독일 등이 높게 나타나 네트워크 내에서 중심 역할을 수행하고 있으며, 매

개 중심성은 미국이 높게 나타나 다른 국가의 연구자들을 연결시켜주는 가교 역할을 하고 있는 것으로 해석할 수 있다.

(표 2) 국가별 중심성 순위

순위	국가	연결중심성	근접중심성	매개중심성
1	United States	0.5909	0.6912	0.3011
2	Canada	0.4773	0.6264	0.1433
3	Germany	0.4545	0.6168	0.1455
4	United Kingdom	0.3636	0.5896	0.0998
5	Netherlands	0.3636	0.5727	0.0647
6	Sweden	0.2955	0.5568	0.0464
7	Australia	0.2727	0.5418	0.0296
8	Finland	0.2500	0.5207	0.0220
9	China	0.2500	0.5275	0.0641
10	Italy	0.2500	0.5275	0.0748

4. 결론

본 연구에서는 최근 활발하게 논의되고 있는 인공지능 윤리와 관련된 국제공동연구 현황을 실증적으로 살펴보기 위해 국가 간 공저논문 기반 공동연구 네트워크 분석을 실시하였다. 아직 인공지능 윤리에 대한 연구가 많이 수행되지는 않았으며 소수의 국가에서 연구가 편중되어 있다. 특히 국가 간 공동연구는 미국을 중심으로 이루어지고 있는 모습을 보이고 있다.

현재 인공지능의 윤리적 개발과 활용에 대해 국제적인 정상회담이 이루어지고 있으며, 국제기구 및 주요 선진국들을 중심으로 인공지능 윤리 준칙에 대한 가이드라인과 법·제도가 만들어지고 있다. 따라서 국가 간 공동연구는 더욱 중요해질 것이며, 활성화하기 노력이 필요하다. 즉, 동 분야의 연구를 주도하고 있는 미국, 영국, 독일, 캐나다를 중심으로, 다양한 연구가 이루어질 수 있도록 범국가적 인공지능 윤리 연구자 네트워크를 구성하여 인공지능 부작용에 국제적으로 대응할 수 있는 체계 마련이 필요하다.

감사의 글

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: KISTI 기본사업 K24L5M1C1-01)

참고 문헌

[1] HAI, “Artificial Intelligence Index Report 2024”, Stanford University, Human-Centered Artificial Intelligence, 2024.
 [2] 유화선, 윤병성, 최희석. “인공지능 윤리(AI Ethics): 인간과 인공지능의 조화로운 공존 방안”, KISTI ISSUE BRIEF, 제68호, 2024.
 [3] 김원진. “과학기술분야 국제협력 증진을 위한 아시아 국가 간 공동연구 현황 분석”, 박사학위논문, 2010.

KDD/KM 방법론을 활용한 미래차 분야 국제협력 연구동향 분석

유화선
한국과학기술정보연구원 정책연구센터
e-mail : hsyoun@kisti.re.kr

The Analysis of International Research Collaboration Trends in the Field of Future Vehicles Using KDD/KM Methodology

Hwasun You
Korea Institute of Science and Technology Information(KISTI)

요 약

최근 글로벌 기술패권경쟁이 심화됨에 따라 국내는 국제협력 연구로의 전환을 추진하고 있다. 본 연구에서는 KDD/KM 방법론을 활용하여 동시출현 키워드 네트워크 분석을 통해 미래차 분야의 국제협력 연구동향을 살펴보았다. 분석 결과, 미국, 중국, 독일의 국제협력 활동이 활발하며, 특히 미국을 중심으로 중국과 독일을 매개로 다양한 국가와의 연구를 수행하고 있는 것으로 분석되었다.

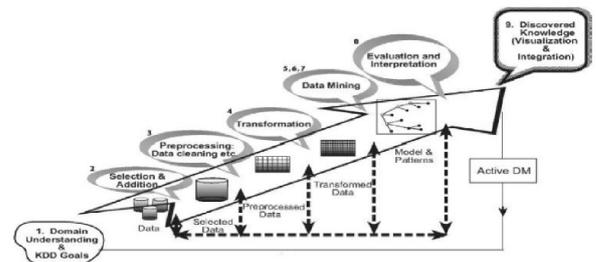
1. 서 론

글로벌 시대 급변하는 사회변화에 대응하기 위해 세계 각국은 다양한 분야에서 국제협력을 활성화하여 전문성을 강화하고 있다. 국제협력은 한정적인 자국의 자원과 인력의 한계를 극복하여 고차원의 연구를 가능하도록 하기 때문에 중요성은 더욱 커지고 있다[1]. 국제협력 연구는 연구자들의 상호교류 방식 중 하나로 문제의식, 연구주제 및 연구방법 등에 대한 생각과 기술을 공유함으로써 연구성과에 영향을 미친다[2]. 이에 따라, 본 연구에서는 최근 국제협력이 강조되고 있는 상황에서 국가 및 기관 간 국제공동연구 현황을 실증적으로 분석하고자 한다.

2. 연구 방법

2.1 KDD/KM 방법론

KDD(Knowledge Discovery in Database) 1996년 Fayyad가 통계적인 패턴 또는 지식을 탐색하는 데 활용할 수 있도록 체계적으로 정리한 프로파일링 기술 기반의 마이닝 프로세스로, 데이터 마이닝, 기계학습, 인공지능, 패턴인식, 데이터 시각화의 구조를 가지고 있다[3]. KM(Knowledge Map)은 데이터로 표시할 경우 사용자가 지식을 찾아내기 어렵기 때문에 데이터 분석 결과를 시각화하여 나타냄으로써 쉽게 지식을 찾아낼 수 있도록 데이터 분석 결과를 시각화시킨 것을 의미한다. KDD/KM 방법론의 분석 절차는 데이터셋 선택, 데이터 전처리, 데이터 변환, 데이터 마이닝, 데이터 마이닝 결과 평가 등의 과정을 거친다[4].



(그림 1) 데이터베이스를 통한 지식발견(KDD) 구성 요소

2.2 분석 방법

본 연구에서는 KDD/KM 방법론을 기반으로 동시출현 키워드 네트워크 분석을 통해 미래차 분야 국제공동 연구 현황을 살펴보고자 하였다. 동시출현 키워드 분석 방식은 일반적으로 방대한 양의 서지사항을 분석하는 방식에 적용되며[5], 공동연구 네트워크 분석은 연구 분야의 경향을 파악하는 연구부터 성과와의 상관성을 살펴보는 연구 등에 많이 활용되고 있다[6].

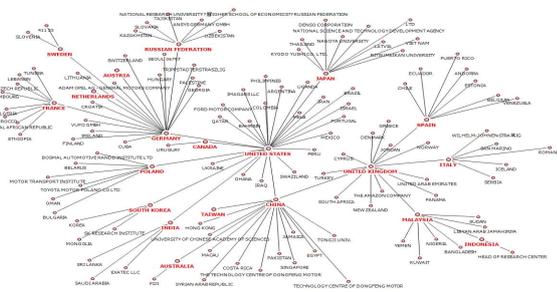


(그림 2) 동시출현 키워드 분석 방법

분석을 위한 데이터는 Scopus에서 발표된 미래차 분야 최근 5년 동안의 SCI 논문 총 16,105건을 활용하였다.

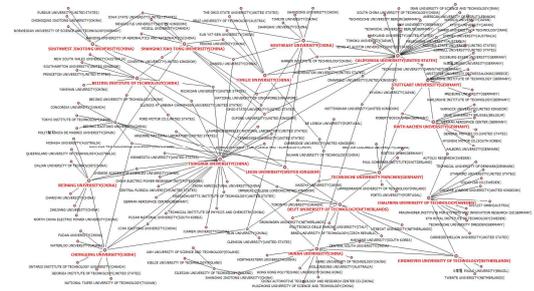
3. 분석 결과

미래차 분야 국제공동연구는 미국, 중국, 독일을 중심으로 협력 네트워크를 형성하고 있다. 미국은 한국뿐만 아니라 인도, 중국, 독일, 폴란드 등과의 국제협력 성과가 높은 반면, 중국은 국제공동연구보다는 국내 연구자들 간의 공동연구 수행이 많은 편인 것으로 나타났다. 한국은 미래차 분야 관련 최근 5년 동안 발표한 513건의 논문 중에서 국제공동연구에 의한 논문은 76건에 불과한 것으로 나타나, 아직까지는 국내 연구자 간 협력 네트워크 위주로 구성되어 있는 것으로 분석되었다. 특히 자율주행차, 전기차, 수소차 등 미래차 분야의 경우, 산업적 특성상 연구자들이 연구 초기 논문보다는 관련 특허 보유를 통해 기술력을 확보하려는 것에 기인한 결과라고 판단된다.



(그림 3) 국가 간 국제협력 네트워크 현황

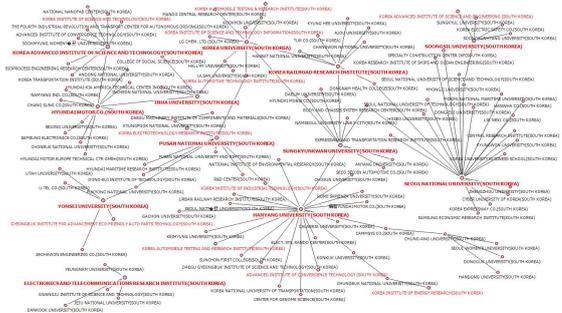
글로벌 기관 간 국제협력 네트워크 분석 결과, 중국의 Tsinghua university 및 Beijing Jiaotong university, 미국의 California university, 독일의 Technische university Munchen가 1차 핵심 그룹을 형성하고 있는 것으로 나타났다. 특히 1차 핵심 그룹 중에서도 미국의 California university가 다양한 공동연구를 수행하고 있는데 1차 그룹 기관과의 협력뿐만 아니라 일본, 호주, 이란 등 다양한 국가의 기관과도 협력 관계를 형성하고 있다.



(그림 4) 기관 간 국제협력 네트워크 현황

중국의 Tsinghua university도 미국에 있는 연구기관과 다양한 공동연구를 수행하고 있는 것으로 나타나며, 이외에도 한국, 독일, 싱가포르 등 국가의 기관과도 공동연구 네트워크를 형성하고 있는 것으로 분석되었다.

다음으로 국내 대학 및 연구기관의 협력 네트워크 분석 결과, 연구활동도 1위인 Seoul National university는 국내 대학과의 공동연구뿐만 아니라 Samsung Economic Research Institute, Hanwha Co., Hyundai Motor Co. 등 산업체와 공동연구를 추진하고 있는 것으로 분석되었으며, 출연연은 Korea Railroad Research Institute, Electronics and Telecommunications Research Institute를 중심으로 협력 네트워크를 형성하고 있는 것으로 나타났다.



(그림 5) 국내 기관 협력 네트워크 현황

4. 결론

최근 전 세계적으로 기술패권경쟁이 심화되면서 국내는 국제협력 연구로의 전환을 추진하고 있다. 특히 국가전략 기술 지정(‘22.10)으로 국제협력연구가 필수인 상황이다. 따라서 국내는 과학기술분야 국제협력을 강화하기 위한 중장기적인 단계별 로드맵 구축이 필요하며, 이를 지속적으로 추진할 수 있는 전략 마련 및 제도 개선이 필요하다.

감사의 글

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: KISTI 기본사업 K24L5M1C1-01)

참고 문헌

- [1] 노영희, 장로사, “한국의 국제연구협력 주제 연구동향에 관한 연구”, 『사회과학연구』, 30(2): 77-92, 2019.
- [2] Friedkin, N. E. “Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study”, *American Sociological Review*, 58(6), 861-872, 1993.
- [3] 나홍석. 이기적 빅데이터분석기사 필기 기본서, 영진닷컴, 2020.
- [4] Umair Shafique, Haseeb Kaiser. “A Comparative Study of Data Mining Process Models”, *International journal of innovation and scientific research*, Vol.20, November 2014.
- [5] 유화선, 김윤명, 양치승. “국내 우수 연구자의 R&D 국제협력 활동도 분석 연구: 가사로봇 분야를 중심으로”, *기술혁신학회지*, 22(4): 605-632, 2019.
- [6] 박연수, 박주영, 김종주, 원유형. “기후변화 학술연구협력 경향 분석: 공동저자 네트워크 구조 및 중개자 특성 분석을 중심으로”, *기술혁신학회지*, 23(1): 162-180, 2020.

정밀 휴먼 디지털 트윈 구현을 위한 비전 트랜스포머 기반 의료영상 초해상화 모델: 무릎 영상 사례

이재용¹, 양동훈¹, 심형섭¹, 황명권^{1,2}, 이경하^{1,2}

¹한국과학기술정보연구원 (KISTI)

²과학기술연합대학원대학교 (UST)

e-mail : {ljy315, yangdonghun3, hsshim, mgh, kyongha}@kisti.re.kr

Vision Transformer-Based Medical Image Super-resolution Model for Developing Precise Human Digital Twin: A Case Study on Knee Images

Jae Yong Lee¹, Donghun Yang¹, Hyung-Seop Sim¹, Myunggwon Hwang^{1,2}, Kyong-Ha Lee^{1,2}

¹Korea Institute of Science and Technology Information (KISTI)

²University of Science and Technology (UST)

요약

정밀 휴먼 디지털 트윈(Human Digital Twin) 구현을 위해서는 고해상도 의료영상 데이터가 필수적이다. 특히 복잡한 무릎 관절 및 연조직 구조 등의 모델을 정밀하게 구축하기 위해서는 고해상도의 이미지가 요구되지만, 기존 의료영상 기술은 고가의 장비, 전문인력, 복잡한 촬영 환경 등 다양한 제약이 따른다. 본 연구는 KISTI의 디지털 코리안(Digital Korean) 데이터셋을 활용하여 기존의 일반 도메인 특화 초해상화 모델인 HAT(Hybrid Attention Transformer) 모델을 의료영상 도메인에 적합하도록 미세 조정하였으며, 이를 무릎 영상 데이터셋에 적용하는 연구를 수행하였다. 적용 결과 기존 모델 대비 PSNR 및 SSIM 지표에서 각각 약 8.5%, 0.9% 향상된 성능을 보였다. 이러한 결과는 제안 모델이 다양한 의료영상 데이터에 효과적으로 적용될 수 있음을 시사하며, 정밀한 휴먼 디지털 트윈 구현에 기여할 수 있음을 보여준다.

1. 서론

휴먼 디지털 트윈은 디지털공간에 정밀한 인체의 복제본을 구축하는 기술로써, 진단·치료 시뮬레이션, 비대면 진료 및 교육훈련 등에 활용될 수 있는 중요한 기술이다[1]. 이러한 휴먼 디지털 트윈을 정밀하게 구현하기 위해서는 CT, MRI 등 고해상도 의료영상 데이터가 요구된다. 특히, 무릎과 같이 복잡한 관절 구조 및 연조직 구조를 갖는 인체 요소를 정밀하게 모델링하기 위해서는 고해상도 의료영상이 필수적이지만, 기존 의료영상 기술은 고가의 장비, 전문인력, 및 복잡한 촬영 환경 등 다양한 제약으로 인해 해상도의 한계를 갖는다[2]. 이러한 문제를 해결하기 위해 최근 딥러닝 기반의 다양한 초해상화 연구들이 수행되고 있으며, 특히 비전 트랜스포머 기반 모델들이 일반 도메인에서 우수한 성능을 보이고 있다[3]. 그러나 이러한 모델들은 일반 도메인 이미지를 중심으로 학습되어, 의료 영상 데이터에 직접 적용하는 데는 한계가 있다.

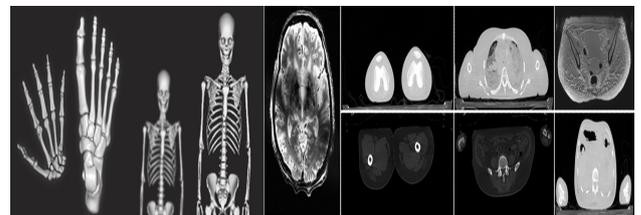
2. 관련 연구

초해상화 기술은 저해상도 이미지를 고해상도로 복원하는 기술로, 선형 보간법 등 전통적인 방법론보다 딥러닝 기반 모델들이 우수한 성능을 보이고 있다[3]. 특히, 비전 트랜스포머 기반 초해상화 모델들은 셀프 어텐션 메커니즘을 중심으로 픽셀 간의 상관관계를 학습하여 보다 효과적으로 고해상도의 이미지를 복원한다. 그러나 대부분의 비전 트랜스포머 기반 모델들은 일반 도메인 이미지를 중심으로 학습되어, 의료 영상과

같은 특수 도메인에서는 초해상화 성능이 저하될 수 있다.

3. 디지털 코리안(Digital Korean) 데이터셋

본 연구는 일반 도메인 이미지 중심으로 사전 학습된 초해상화 모델을 미세 조정하기 위해 KISTI의 디지털 코리안 데이터셋을 활용하였다[4]. 그림 1과 같이 해당 데이터셋은 한국인의 인체 골격 표준 모델을 개발하기 위해 20세에서 60세 사이의 한국인 남녀 100구의 시신을 대상으로 1mm 이하 간격의 CT 스캔을 통해 구축되었으며, 총 512x512 해상도 이미지 195,622장으로 구성된다.



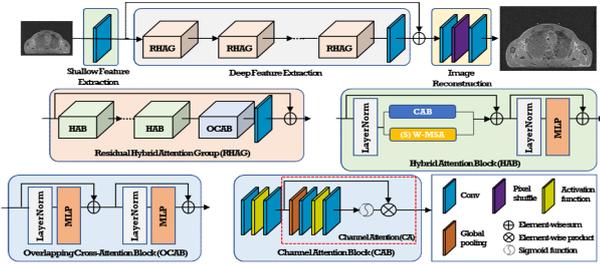
(그림 1) 디지털 코리안 데이터셋 예시[4]

4. 비전 트랜스포머 기반 의료영상 특화 초해상화 모델

4.1 일반 도메인 사전학습 모델

본 연구에서 활용한 일반도메인 특화 사전학습 초해상화 모델은 비전 트랜스포머 기반의 HAT 모델이다. HAT 모델은 채널 어텐션과 셀프 어텐션을 결합해 픽셀 간 정보를 효과적

으로 통합할 수 있으며, 그림 2와 같이 얇은특징 추출, 깊은특징 추출, 이미지 재구성 세 단계로 구성된다[5]. 또한 L1 손실 함수를 통해 고해상도 이미지를 정밀하게 복원할 수 있으며, 일반 도메인에서 우수한 초해상화 성능을 보인다.



(그림 2) HAT 모델 구조[5]

4.2 미세 조정

본 연구는 일반 도메인 영상을 중심으로 학습된 기존의 HAT 모델을 디지털 코리안 데이터셋으로 미세 조정된 비전 트랜스포머 기반 의료영상 특화 초해상화 모델을 제안한다. 미세조정 절차는 다음과 같다. 첫째, HAT 모델 구조를 백본으로 정의한다. 둘째, HAT 연구에서 제공하는 일반 도메인의 이미지들로 사전 학습된 모델 가중치를 이용하여 HAT 모델의 가중치들을 초기화한다. 셋째, 디지털 코리안 데이터셋을 기반으로 미세조정한다.

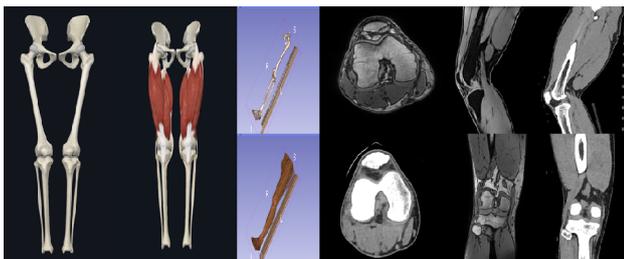
5. 실험

5.1 실험 환경

본 연구에서 제안한 비전 트랜스포머 기반 의료영상 특화 초해상화 모델은 Python 3.8과 PyTorch 1.13.1 환경에서 NVIDIA A100 (80GB)*2를 활용해 구현되었다. 배치 크기는 2로 설정하였으며, 이외 모든 하이퍼 파라미터는 기존 HAT 모델과 동일한 파라미터 값을 사용하였다. 디지털 코리안 데이터셋 195,622장을 활용하여 기존의 HAT 모델을 미세조정 하였으며, 제안 모델의 성능 검증을 위해서는 무릎 의료영상 데이터 5,000장을 활용하였다. 성능 평가 지표로는 PSNR(최대신호 대 잡음비) 및 SSIM(구조적 유사도)를 사용하였다.

5.2 무릎 의료영상 데이터셋

본 연구는 제안모델의 성능 검증을 위해 그림 3과 같이 복잡한 관절 구조 및 연조직 구조를 갖는 무릎에 대한 총 147,831장(40세트)의 CT/MRI 데이터를 구축하였다.



(그림 3) 무릎 의료영상 데이터셋 예시

5.2 실험 결과

표 1은 일반 도메인 이미지 데이터셋만으로 사전 학습된 기존 HAT 모델과 본 연구에서 제안하는 디지털 코리안 데이터셋으로 미세 조정된 비전 트랜스포머 기반 의료영상 특화 초해상화 모델의 무릎 의료영상 초해상화 성능을 보인다. 본 연구의 제안 모델이 PSNR 기준 40.8295, SSIM 기준 0.9719로, 기존 모델 대비 약 8.5%(PSNR), 0.9%(SSIM) 우수한 성능을 보였다.

(표 1) 정량평가 성능 비교

	기존 HAT 모델	제안 모델
PSNR	37.6470	40.8295
SSIM	0.9628	0.9719

6. 결론 및 향후연구

본 연구는 정밀한 휴먼 디지털 트윈 구현을 위해 필수적인 고해상도 의료영상 데이터를 구축하고자, 기존 일반 도메인 특화 초해상화 모델(HAT)을 KISTI의 디지털 코리안 데이터셋으로 미세 조정된 비전 트랜스포머 기반 의료영상 특화 초해상화 모델을 제안하였으며, 기존 HAT 모델 대비 우수한 초해상화 성능을 보였다. 이러한 결과는 제안 모델이 다양한 의료영상 데이터에 효과적으로 적용될 수 있음을 시사하며, 정밀한 휴먼 디지털 트윈 구현에 기여할 수 있음을 보여준다. 하지만 본 연구의 결과는 오직 무릎 데이터셋에 대한 결과이다. 따라서 추후 다양한 인체 요소 영상 데이터에 본 연구의 제안모델을 적용하여 성능을 추가 검증할 예정이다.

Acknowledgement

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본 사업으로 수행된 연구입니다. (과제번호: (KISTI)K24L1M2C3)

참고 문헌

- [1] 이경하, 박정현, 양동현, 심형섭, 이재용. “휴먼 디지털 트윈: 인체 연구에 대한 효과적 대체제”, 『KISTI 이슈브리프』 제64호, 2023.
- [2] 이재용, 양동현, 이경하. “정밀한 휴먼 디지털 트윈 구현을 위한 의료영상의 초해상화”, 『Korean DataBase Conference (KDBC) 2023 논문집』 pp.38-41, 2023.
- [3] 이재용, 양동현, 박정현, 심형섭, 이경하. “휴먼 디지털 트윈을 위한 초해상화 관련 기술 동향”, 『대한산업공학회 춘계공동학술대회 논문집』 pp.4449-4460, 2023.
- [4] 최광남, 한승호. “디지털 코리안 인체 모델 구축”, 『지식정보인프라』 제21권, pp. 20-25, 2006.
- [5] X. Chen, X. Wang, J. Zhou, Y. Qiao, C. Dong, “Activating more pixels in image super-resolution transformer. in Proc. IEEE/CVF International Conference on Computer Vision, Paris, France, 2023, pp. 22367-22377.

페르소나를 이용한 인공지능 문헌 검토자의 적절성 연구

이태석, 선충녕, 한상준, 유수현
한국과학기술정보연구원 융합서비스센터
tsyi@kisti.re.kr, wilowisp@kisti.re.kr, sjhan@kisti.re.kr, yoosu@kisti.re.kr

Using personas to study the appropriateness of AI literature reviewers

Taeseok Lee, Choong-Nyoung Seon, Sangjun Han, Suhyeon Yoo
KISTI Convergence Service Center

요약

연구자들은 선행연구를 정리하거나, 관련지식을 습득하기 위해 문헌 조사를 수행한다. 또한 새로운 연구를 준비하기 위해 연구 동향을 파악하고, 연구 주제를 탐색하며, 연구 방법을 모색하는 것이 필요하다. 문헌 조사는 연구 초기 단계에 지식을 정리하거나 연구를 준비할 때, 자주 활용하는 도구이다. 최근, 거대언어모델의 발전에 따라, 이러한 문헌 조사를 수행하는 AI 논문 검토자를 생성하여 문헌 조사 과정에 도움을 줄 수 있는 방법을 제시하고자 한다. 본 논문은 거대언어모델을 사용하여 여러 논문의 핵심 관점을 추출하고, 비교 분석한 정보로 AI 논문 검토자의 페르소나를 생성하고, 가상의 대화자의 적합성을 평가하였다. 그 결과, 다양한 AI 검토자의 페르소나를 생성하고 대화할 수 있으며, 이를 통해 문헌 조사의 목적을 효과적으로 달성할 수 있음을 확인하였다.

1. 서론

연구자들은 선행연구를 정리하거나, 관련지식을 습득하기 위해 문헌 조사를 수행한다. 또한 새로운 연구를 준비하기 위해 연구 동향을 파악하고, 연구 주제를 탐색하며, 연구 방법을 모색하는 것이 필요하다. 문헌 조사는 연구 초기 단계에 지식을 정리하거나 연구를 준비할 때, 자주 활용하는 도구이다.

거대언어모델(LLM, Large Language Model)의 발전과 대화형 AI에 대한 연구를 통해 페르소나 정보를 활용한 보다 나은 대화 응답 생성이 가능해졌다[1][2]. 이를 바탕으로 문헌 조사를 지원하는 AI 검토자로서의 활용 가능성을 탐색하고, 문헌 조사 수행에 필요한 지식을 효율적으로 정리하는 데 도움을 받을 수 있는지 확인하고자 한다. 이를 위해서는 AI 문헌 검토자의 성격과 특성을 결정할 수 있는 적절한 페르소나를 생성하는 것이 필요하다.

허깅페이스에 공개되는 LLM의 크기는 2B에서 405B까지 다양화되었으며, 양자화 기술의 적용으로 더욱 선택의 폭이 더욱 넓어졌다. 이들 모델 중에서 문헌조사를 수행할 수 있는 적합한 모델을 선택할 필요가 있으며, 문헌 조사를 수행하는 AI 논문 검토자를 생성하기 위한 프롬프트와 대화형 UI를 개발해야 한다.

본 논문에서 제안한 방법은 여러 논문의 핵심 관점을 추출하고, 이를 비교 분석하여 AI 논문 검토자의 페르소나를 생성하며, 가상의 대화자의 적합성을 평가하는 것이다. 그 결과, 다양한 AI 검토자의 페르소나를 생성하고 대화할 수 있으며, 이를 통해 문헌 조사의 목적을 효과적으로 달성할 수 있음을 확인하였다.

2. 문헌조사에 대한 이해

2.1 문헌조사의 정의 및 목적

문헌 조사는 연구 중인 주제와 관련된 선행 연구를 객관적이고 간결하며 비판적으로 요약한 것이다. 또한, 특정 주제나 연구 질문과 관련된 이용 가능한 문헌, 연구 및 학술 간행물을 종합적으로 분석하는 과정이라고 할 수 있다.

문헌조사는 연구자와 학자가 자신의 분야에서 선행연구를 정리하거나, 관련지식을 발전시키고자 할 때 활용된다. 또한 새로운 연구를 준비하기 위해 연구 동향 파악, 연구 주제 탐색, 연구 방법 탐색 과정이 필요하다. 문헌 조사는 연구 초기 단계에 지식을 정리하거나 연구를 준비하는 중요한 도구이다. 따라서 문헌조사의 목적은 선행 연구 정리, 관련 지식 습득, 연구 동향 파악, 연구 주제 탐색, 그리고 연구 방법 탐색이라고 할 수 있다[3].

2.1 문헌조사의 필요성

문헌 검토는 연구의 맥락을 설정하는 데 도움을 주며, 연구 목표의 틀을 구성하고, 연구 질문을 개발하고, 연구 방법론을 설계하는 데 기여한다. 또한, 이전 연구에서 사용된 관련 이론, 방법론 및 접근 방식을 식별하고 향후 연구를 구체화하는 데 유용하다. 이처럼 문헌조사는 연구 주제를 선정하는 과정에서 중요한 역할을 한다[3].

2.3 문헌조사의 기능

해당 분야의 기존 지식에 대한 개요를 제시하고, 격차 및 불일치를 식별하며, 추가 조사가 필요한 영역을 명확히 한다. 또한, 특정 분야 또는 특정 질문에 대한 기존 연구

를 요약하고 평가하며 비판적으로 분석하는 작업을 수행한다.

이러한 기능에 따라 논문에서 추출해야 하는 핵심 관점을 정의하였다. 핵심 관점은 다음과 같이 11개로 식별되었다: “연구 내용 요약”, “연구 목적”, “연구의 제약 및 한계점”, “핵심 요점, 주장 및 근거”, “선행 연구와의 차별성”, “논문의 영향력 및 가치”, “논리의 공백이나 검증이 필요한 부분”, “향후 연구”, “연구 방법 및 도구”, “연구 자금 지원”, “연구 기간.” 또한, 조사 목적에 따라 논문의 주요 관점을 표 1과 같이 맵핑하였다.

(표 1) 조사목적에 따른 논문의 주요관점

조사목적	논문의 주요관점
선행 연구 정리	연구 목적, 선행 연구와 차별성, 연구 방법 및 도구
관련 지식 습득	연구 내용 요약, 논문의 영향력 및 가치, 연구 방법 및 도구
연구 동향 파악	향후 연구, 연구 방법 및 도구, 연구 자금 지원
연구 주제 탐색	연구의 제약 및 한계점, 논리의 공백이나 검증이 필요한 부분, 향후 연구
연구 방법 탐색	연구 목적, 연구 방법 및 도구, 연구 기간

3. AI 문헌조사 구현 및 페르소나 평가

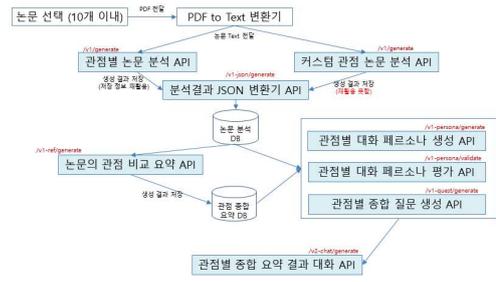
AI 문헌조사 UI는 그림 1과 같이 구성하였다. 각 기능의 실행을 위해서 그림 2와 같이 기능 API를 구현하였다. LLM 실행기로 vLLM을 사용하고, API는 FastAPI로 구현했으며, UI는 gradio를 활용하였다. 모델의 선택은 논문 PDF를 텍스트 변환했을 때, 20쪽 분량(32KB)을 입력할 수 있는 모델 중에서 이루어졌다. 또한, 논문 관점 추출, 대화, 지시문의 맥락 이해도, 한국어 생성 등의 성능을 고려하였다. Mistral 7B와 Qwen2 7B는 영어 생성에서는 우수한 성능을 보였으나, 한국어 생성에서는 불안정한 결과를 나타냈다. 한국어 생성이 안정적인 모델은 Mistral-nemo와 Llama 3.1 모델이었으며, 성능 면에서는 Mistral-nemo가 우수했으나, Llama 3.1은 모델 크기가 작고 생성 속도가 더 빨라 최종적으로 선택하였다.



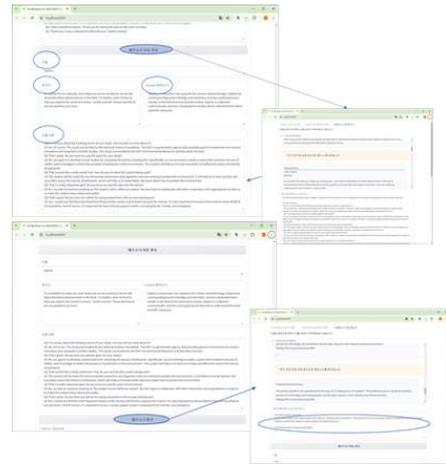
(그림 1) AI 검토자 대화 UI

AI 검토자 페르소나는 그림 2와 같이 논문의 주요 관점과 비교하여 요약한 내용을 입력하여 이름, 첫인사, 성격(이력), 선행 대화를 생성하고, 그 결과를 LLM을 통하여 10점 만점으로 평가하는 과정을 그림 3과 같이 수행하였다. 그 결과 평균 8점이 나타났으며, 필요한 경우 수정하

여 저장한 후 대화에 활용하였다.



(그림 2) AI 문헌조사 기능 API 관계도



(그림 3) AI 검토자 페르소나 생성 및 평가

4. 결론

논문 문헌조사에 대한 정의와 필요성, 그리고 기능을 살펴본 후, 조사 목적에 따른 논문 분석의 주요 관점을 도출하였다. LLM을 활용하여 AI 검토자를 생성하고 주요 관점에 따라 대화하는 것이 가능하였다. AI 검토자의 페르소나를 적절히 설정함으로써 보다 다양한 관점에서 대화가 가능해졌다.

AI 검토자의 페르소나 적합성 평가를 LLM을 사용하여 진행하였으나, 보다 객관적인 평가를 위한 후속 연구가 필요하다.

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다(과제번호:K24L1M3C4).

참고 문헌

[1] Liu, J., Symons, C., & Vatsavai, R. “Persona-Based Conversational AI: State of the Art and Challenges”, 2022 IEEE International Conference on Data Mining Workshops (ICDMW), 2022. pp.993-1001.

[2] Wang, Y., Xu, F., Wang, R., Zhu, Z., Liu, P., & Lu, R. “Improving Persona Understanding for Persona-based Dialogue Generation with Diverse Knowledge Selection”, 2022 26th International Conference on Pattern Recognition, 2022. pp.1915-1921.

[3] Snyder, H. “Literature review as a research methodology: An overview and guidelines”, Journal of Business Research, Vol. 104, 2019. pp.333-339.

흐린 이미지 재구성을 위한 3D 가우시안 스플래팅 향상

원일남¹, 조인휘²

¹한양대학교 컴퓨터소프트웨어학과 석사과정

²한양대학교 컴퓨터소프트웨어학과 교수

ynyuan@hanyang.ac.kr, iwjoe@hanyang.ac.kr

Improving 3D Gaussian Splatting for Blurred Image Reconstruction

Yinan Yuan¹, Inwhee Joe²

¹Dept. of Computer Science, Hanyang University

²Dept. of Computer Science, Hanyang University

요약

본 연구는 3D 장면 재구성과 새로운 뷰 합성을 위해 3D 가우시안 스플래팅(3DGS)을 활용하는 방법을 제안한다. 3DGS는 우수한 합성 품질로 주목받고 있으나, 훈련 이미지의 블러가 장면 재구성 품질을 저하시킬 수 있으며, 많은 훈련 뷰가 필요하다는 한계가 있다. 이를 해결하기 위해, 본 연구는 희소한 각도의 블러 입력으로부터 명확한 가우시안 분포를 복원하는 방법을 제안한다. 소형 MLP를 사용하여 3D 가우시안의 공분산을 조정함으로써 장면의 블러를 모델링하고, 다양한 손실 함수와 최적화 전략을 통해 고품질 3D 뷰를 생성한다. 이 방법은 Deblur-NeRF 데이터셋에서 우수한 성능을 보였다.

1. 서론

최근 3D 재구성 분야는 신경 방사 필드(NeRF)와 3D 가우시안 스플래팅(3DGS) 같은 새로운 기술의 등장으로 크게 발전하였다. NeRF는 고품질의 3D 장면 합성을 가능하게 하였지만, 계산 복잡성과 긴 훈련 시간으로 인해 한계점도 드러냈다. 이에 반해 3DGS는 점구름을 가우시안 분포로 표현함으로써 높은 계산 효율성과 희소한 데이터에 대한 강한 강건성을 제공하지만, 복잡한 장면에서는 흐림 효과가 나타날 수 있다. 본 연구는 3DGS의 가우시안 분포 공분산을 미세 조정하여 흐림 효과를 감소시키고, 희소한 입력 데이터로부터 고품질의 3D 재구성을 달성하는 방법을 제안한다. 이 방법은 3D 재구성 분야에서 더 효율적이고 신뢰할 수 있는 솔루션을 제시할 것이다.

2. 본론

2.1 블러 이미지 학습

본 연구에서는 3D 재구성 작업에서 블러 처리를 효과적으로 수행하기 위해 다층 퍼셉트론(MLP)과 다양한 손실 함수를 결합한 모델을 제안하였다. 기존 연구에 따르면, 3D 재구성 작업에서 블러된 이미지를 먼저 학습한 후 원래 매개변수를 조정하는 방법이 효과적임을 알 수 있다.[1] MLP는 신경망의 가장 기본적인 형태로, 구조가 단순하고 구현과 훈련이 용이하며 비선형 관계를 처리하는데 뛰어난 성능을 보인다. MLP는 다층 구조와 비선형 활성화 함수를 통해 복잡한 패턴과 관계를 포착할 수 있어, 이미지 블러 처리와 가우시안 분포의 공분산 행렬 조정에 중요한 역할을 한다. 또한, MLP는 상대적으로 데이터 요

구량이 낮아 제한된 데이터로도 효과적인 훈련이 가능하며, 3D 가우시안 분포의 공분산 행렬을 조정하는 데 충분한 표현력과 학습 능력을 제공한다. 본 연구는 초점 블러와 카메라 움직임 블러라는 두 가지 주요 블러 유형을 처리하기 위해 MLP를 활용하여, 3D 가우시안 분포의 위치, 회전, 스케일링 요소를 조정함으로써 고품질의 3D 재구성을 가능하게 한다. 이 방법은 MLP를 통해 카메라의 위치와 회전 상태를 시간에 따라 보간하여 카메라 움직임을 시뮬레이션하고, 학습된 조정을 사용해 명확한 이미지를 생성함으로써 모델의 효율성을 크게 향상시킨다.

2.2 다중 손실 함수 최적화

Pearson Loss와 국부 Pearson Loss를 도입하여 깊이 맵의 전역적 일관성과 로컬 디테일을 최적화했으며, Score Distillation Sampling(SDS) 손실을 통해 노이즈 확산과 지식 증류를 결합하여 모델이 더 자연스럽게 매끄러운 이미지를 생성할 수 있도록 하였다. Alpha Blending Depth loss는 복잡한 장면에서도 정확한 깊이 맵을 생성하는 데 기여한다.

2.3 플로터 제거

생성된 이미지가 전반적인 손실 함수 처리를 거친 후에도 일부 이미지에서는 부유하는 물체로 인해 그림자가 관찰될 수 있다. 이를 해결하기 위해 플로터 제거(Floater Pruning)라는 최적화 기법이 사용된다.[2] 이 기법은 각 가우시안 포인트가 이미지 품질에 기여하는 정도를 평가하여, 중요도가 낮은 가우시안 포인트는 제거하여 모델의

효율성과 성능을 향상시킨다. 본 연구에서는 사진 학습된 깊이 추정 모델인 BoostingMonocularDepth를 [3] 활용하여 가짜 진실 깊이 맵을 생성하고, 가우시안 분포로부터 생성된 깊이 맵과의 상관성을 피어슨 상관 계수로 계산하였다. 이를 통해 스케일 및 오프셋 문제를 방지할 수 있다. 렌더링 과정에서는 알파 블렌딩 깊이 맵과 국부 피어슨 깊이 맵을 고려하여 광선 방향에 따른 모든 가우시안 분포와 픽셀에 가장 큰 기여를 하는 가우시안 분포를 각각 반영한다. Score Distillation Sampling 기법을 통해 생성 모델이 새로운 카메라 뷰를 생성하며, 이러한 뷰에서 부유하는 물체의 영향을 계산하여 최적화 및 제거를 수행한다. 이러한 플로터 제거 기술을 통합함으로써, 본 연구는 모델의 계산 및 저장 요구 사항을 줄이면서도 생성된 이미지의 품질을 유지하고, 모델의 효율성과 성능을 향상시킬 수 있었다.

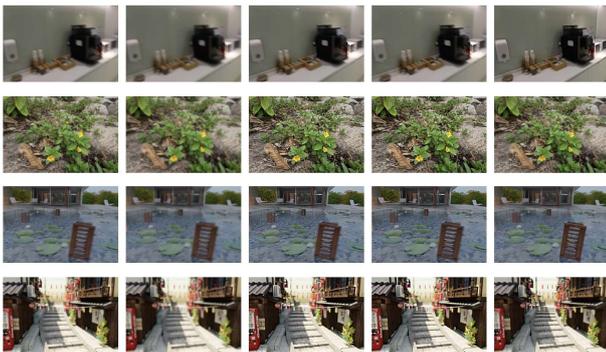
3. 실험

3.1 데이터셋

실험에서 사용된 Deblur-NeRF [1] 데이터셋은 카메라 움직임 블러와 초점 블러가 포함된 실제 이미지와 합성 이미지를 포함하고 있다. 이 데이터셋은 각 장면별로 서로 다른 시점에서 촬영된 고해상도 RGB 이미지 세트를 제공하며, 각 RGB 이미지에 대응하는 고정밀 깊이 맵과 카메라 매개변수도 함께 제공한다.

3.2 실험결과

본 연구에서는 Deblur-NeRF 데이터셋을 기반으로 SV-DRGS 방법과 3DGS [4], NeRF [5], 그리고 Deblur-NeRF 방법을 비교 분석하였다.



(그림 1) (블러 제거 이미지 생성)

(표 1) (정량적 결과 비교)

Model	Real Motion Blur	Real Defocus Blur	Synthetic Motion Blur	Synthetic Defocus Blur
	PSNR SSIM LPIPS	PSNR SSIM LPIPS	PSNR SSIM LPIPS	PSNR SSIM LPIPS
3DGS	21.874 0.644 0.374	21.444 0.620 0.356	22.404 0.622 0.324	24.626 0.736 0.232
NeRF	23.488 0.702 0.330	23.156 0.678 0.330	24.270 0.694 0.336	25.806 0.778 0.228
Deblur-NeRF	25.902 0.788 0.234	27.286 0.794 0.222	28.406 0.856 0.236	28.346 0.856 0.118
Ours	27.448 0.822 0.112	28.070 0.850 0.126	28.224 0.860 0.110	28.346 0.896 0.102

그림 1에서 보는 바와 같이 첫 번째 열은 원본 입력 이미지이고, 나머지 열은 3DGS, NeRF, Deblur-NeRF 및 본 연구에서 제안한 방법으로 처리된 생성 이미지이다. 우리는 실제 카메라 움직임 흐림, 실제 초점 흐림, 합성된 카메라 움직임 흐림, 합성된 초점 흐림이 포함된 네 부분의 데이터셋에서 각각 5개의 장면을 선정하여 실험을 수행하였다. 이 다섯 장면의 평가 지표 평균값을 결과로 사용하였으며, 표 1에서 볼 수 있듯이, 본 연구는 모든 장면에서 PSNR, SSIM, LPIPS 지표에서 3DGS, NeRF, Deblur-NeRF보다 우수한 결과를 보였다.

4. 결론

본 연구에서는 가우시안 스플래팅 기법을 기반으로 한 디블러링 재구성 방법인 SV-DRGS를 제안하였다. 이 방법을 통해 희소한 뷰에서도 고품질의 디블러링된 360도 파노라마 이미지를 생성할 수 있다. MLP와 다양한 손실 함수를 결합함으로써 뷰 일관성과 깊이 정확성을 크게 개선하였다. Deblur-NeRF 데이터셋을 활용한 실험 결과, SV-DRGS는 카메라 모션 블러 및 초점 블러 이미지에서 기존의 3DGS, NeRF, Deblur-NeRF 모델보다 우수한 성능을 나타내었다. 또한, 플로터 제거 기법을 도입하여 모델의 계산 효율성과 저장 요구사항을 최적화하였다. 향후 연구에서는 모델의 학습 과정 최적화, 더 많은 응용 시나리오 탐색 및 대규모 데이터셋 검증에 집중할 것이며, 추가적인 연구와 최적화를 통해 SV-DRGS가 실제 환경에서 중요한 역할을 할 것으로 기대된다.

참고 문헌

[1] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V. Sander. Deblur-nerf: Neural radiance fields from blurry images. arXiv preprint arXiv:2111.14292, 2021.

[2] Zongxin Ye, Wenyu Li, Sidun Liu, Peng Qiao, and Yong Dou. Absgs: Recovering fine details for 3d gaussian splatting. arXiv preprint arXiv:2404.10484, 2024.

[3] Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient high-quality compact surface reconstruction in unbounded scenes. arXiv:2404.10772, 2024.

[4] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics, 42(4), July 2023.

[5] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In ECCV, 2020.

스타일 전환이 가능한 3D 객체 재구성

권순양¹, 조인휘²

¹한양대학교 컴퓨터소프트웨어학과 석사과정

²한양대학교 컴퓨터소프트웨어학과 교수

k2023184794@hanyang.ac.kr, iwjoe@hanyang.ac.kr

3D Object Reconstruction with Style Transfer

Chunyang Quan¹, Inwhee Joe²

¹Dept. of Computer Science, Hanyang University

²Dept. of Computer Science, Hanyang University

요약

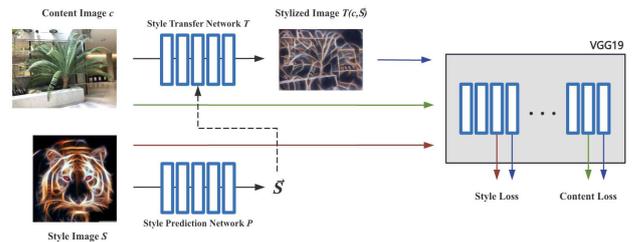
신경 방사 필드(NeRF)는 2D 이미지에서 사실적인 3D 장면을 생성하는 데 탁월한 성능을 보여주 고 있지만, 스타일 전이를 3D 장면에 적용하는 연구는 부족한 실정입니다. 본 논문에서는 NeRF와 스타일 전이를 결합한 새로운 프레임워크를 제안하여, 밀도 추정 없이도 스타일화된 3D 장면을 생성할 수 있도록 합니다. 이를 위해 NeRF의 기하학적 및 외관적 표현 능력을 활용하며, 고급 스타일 전이 기술로 장면의 미적 속성을 수정합니다. 이 방법은 NeRF 모델을 단순화하여 3채널 이미지에 효과적으로 작동하도록 하였고, 실험 결과는 예술적 콘텐츠 제작 및 가상 현실과 같은 다양한 분야에서의 잠재력을 보여줍니다.

1. 서론

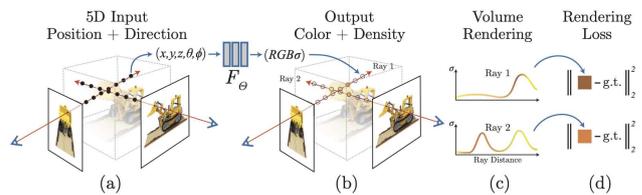
3D 장면 재구성과 렌더링은 컴퓨터 그래픽스와 비전 분야에서 중요한 주제로, 특히 Neural Radiance Fields (NeRF) [1]는 2D 이미지에서 사실적인 3D 장면을 생성하는 데 탁월한 성능을 보였습니다. 하지만 기존 연구는 주로 사실적인 표현에 초점을 맞추었고, 예술적 스타일 전환에 대한 연구는 부족했습니다. 본 논문에서는 NeRF와 스타일 전이 기법을 결합하여, 밀도 추정을 생략하고 RGB 색상 정보만을 활용함으로써 3D 장면의 기하학적 구조를 유지하면서도 예술적인 스타일화를 실현하는 새로운 프레임워크를 제안합니다. 제안된 방법은 가상 현실, 예술 창작 등 다양한 응용 분야에서 창의적이고 예술적인 가능성을 열어줍니다.

2. 본론

본 논문에서는 스타일 전이와 수정된 NeRF 모델을 결합하여 스타일화된 3D 장면을 생성하는 구체적인 방법을 제안합니다. 스타일 전이 부분에서는 먼저 사전 훈련된 컨볼루션 신경망 (VGG19 [2])을 사용하여 콘텐츠 이미지와 스타일 이미지에서 중간 레이어 특징을 추출하고, 그래프 행렬을 계산하여 스타일 이미지의 텍스처와 색상 정보를 포착합니다 [3]. 그런 다음, 콘텐츠 손실과 스타일 손실을 결합한 최적화된 손실 함수를 통해 원본 콘텐츠 구조를 유지하면서도 목표 스타일을 정확하게 반영하는 이미지를 생성합니다. 생성된 스타일화된 이미지는 여러 시점에서 캡처된 깊이 정보와 결합되며, 이러한 결합된 데이터는 NeRF 모델에 입력됩니다.



(그림 1) (스타일 전환 파이프라인)



(그림 2) (NeRF 프레임워크 [1])

NeRF 부분에서는 전통적인 NeRF 모델을 수정하여 불륨 밀도 추정을 제거하고, RGB 색상 값만을 출력하도록 하여 계산 복잡도를 줄였습니다. 이러한 단순화는 3D 공간에서 색상 정보를 효과적으로 처리할 수 있게 하였으며, 특히 스타일화된 이미지와 깊이 정보의 결합에 중점을 둡니다. 이렇게 생성된 스타일화된 이미지와 정확한 깊이 정보를 함께 NeRF 모델에 입력함으로써 3D 기하학적 구조를 손상시키지 않으면서 스타일 변환을 실현할 수 있습니다.

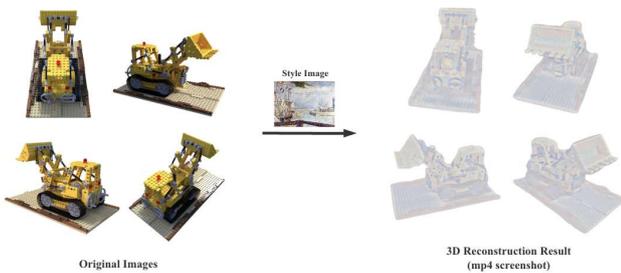
다. 렌더링 과정에서는 볼륨 렌더링 기술을 사용하여 각 카메라 시점에서 투사된 광선 경로의 샘플링된 색상 값을 통합하여 최종적으로 스타일화된 3D 이미지를 생성합니다. 이 방법은 장면의 기하학적 정밀도를 유지하면서도 복잡하고 정교한 스타일 변환을 구현하여 가상 현실, 예술 창작, 그리고 인터랙티브 미디어 분야에서 새로운 가능성을 제시합니다.

3. 실험결과

본 연구에서는 제안된 프레임워크의 성능을 평가하기 위해 다양한 실험을 수행하였습니다. 실험은 두 가지 주요 데이터셋을 사용하여 진행되었으며, 각각은 실제 장면과 합성 장면을 포함합니다.



(그림 3) (3D 재구성 결과1)



(그림 4) (3D 재구성 결과2)

첫 번째 실험에서는, 제안된 NeRF 모델과 스타일 전이 기술을 결합한 방법의 성능을 평가하기 위해 실제 장면 데이터셋을 사용하였습니다. 스타일화된 이미지와 원본 이미지의 깊이 정보를 함께 NeRF 모델에 입력하여 3D 장면을 재구성한 결과, 제안된 방법이 기하학적 구조를 유지하면서도 스타일 정보가 정확하게 반영된 3D 장면을 생성할 수 있음을 확인하였습니다. 특히, 다양한 각도에서 활

영한 이미지에서도 일관된 스타일화 결과를 얻을 수 있었습니다.

두 번째 실험에서는 합성된 데이터셋을 사용하여, 제안된 모델의 스타일 전이 성능을 평가하였습니다. 합성 데이터셋에서는 다양한 조명 조건과 복잡한 장면 구조가 포함되어 있으며, 제안된 방법은 이러한 조건에서도 높은 품질의 스타일화된 3D 장면을 성공적으로 생성하였습니다. 또한, 기존의 NeRF 모델과 비교했을 때, 제안된 방법이 더 적은 계산 비용으로도 유사한 수준의 이미지 품질을 제공할 수 있음을 확인하였습니다.

4. 결론

본 논문에서는 Neural Radiance Fields (NeRF)와 스타일 전이 기술을 결합하여 3D 장면의 예술적 스타일화를 실현하는 새로운 프레임워크를 제안하였습니다. 제안된 방법은 NeRF 모델을 단순화하여 밀도 추정을 생략하고 RGB 색상 정보만을 활용함으로써, 스타일 전이와의 통합을 용이하게 하였습니다. 실험 결과, 제안된 방법은 다양한 데이터셋에서 기하학적 구조를 유지하면서도 고품질의 스타일화된 3D 장면을 생성할 수 있음을 입증하였습니다. 이 연구는 가상 현실, 예술 창작 및 기타 응용 분야에서 새로운 가능성을 제시하며, 2D 스타일 전이와 3D 장면 합성 간의 경계를 허물고 NeRF 기술의 창의적이고 예술적인 응용을 확대할 수 있는 발판을 마련하였습니다. 앞으로의 연구에서는 더욱 복잡한 장면이나 동적인 장면에도 본 방법을 적용하여 실시간으로 스타일화된 3D 장면을 생성할 수 있는 가능성을 탐구하고자 합니다.

참고문헌

[1] BenMildenhall, PratulP.Srinivasan, MatthewTancik, JonathanT.Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In ECCV, 2020.

[2] Karen Simonyan and Andrew Zisserman. Very deep convolutional net-works for large-scale image recognition, 2015.

[3] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style, 2015.

LDA 모델 기반 디지털 트윈 연구 트렌드 분석

고미현¹⁾, 이수철²⁾

1) 한국과학기술정보연구원 지능형시뮬레이션센터,

2) 상명대학교 경영학부(교신저자)

e-mail : 1) mihyungo@kisti.re.kr, 2) suchul.lee@smu.ac.kr

Analysis of Research Trends in Digital Twin Based on LDA Model

Mihyun Ko¹⁾, Suchul Lee²⁾

1) Korea Institute of Science and Technology Information, Intelligent Simulation Center,

2) Department of Business Administration, Sangmyung University(Corresponding Author)

요약

제조 분야를 중심으로 경쟁력 강화와 기술 우위 확보를 통한 시장 선점을 위해 디지털 트윈(digital twin) 기술이 빠르게 적용되고 있다. 최근에는 국민 생활과 밀접한 도시와 안전분야, 환경 이외에도 의료, 에너지, 농·축·수산 등 다양한 공공분야로 디지털 트윈의 적용 분야가 확장하고 있는 추세이다. 이와 같이 다양한 응용 분야에 디지털 트윈 기술이 적용되면서 산업 분야의 효율성 및 생산성 향상, 사회 안전, 국민 삶의 질 향상 및 디지털 기반 고부가가치 일자리 창출 등을 다양한 효과가 나타날 것으로 기대하고 있다.

주요국들은 디지털 트윈 등 제조혁신 기술을 선도하기 위해 공격적으로 관련 정책 및 투자를 확대하고 있으며 국내에서도 최근의 ‘신 디지털 제조혁신 추진전략(2023.09)’ 등 노력을 이어나가고 있다. 하지만 국내 기술 수준은 최고기술을 보유하고 있는 미국에 비해 82.3%로 1.4년의 격차를 보이는 등(정보통신기획평가원, 2021) 디지털 트윈 기술의 경쟁력이 다소 취약한 상황으로 장기적 관점의 기술 내재화 전략이 필요한 시점이다.

이에, 본 연구에서는 토픽 모델링 기법을 활용하여 디지털 트윈의 최신 연구 동향과 특성을 파악하고자 한다. 분석을 위해 WoS (Web of Science)를 통해 관련 논문 데이터를 수집하고 디지털 트윈 연구의 일반적인 특성 파악 및 연구 토픽, 연구 키워드 분석을 수행하고자 한다. 이를 통해, 디지털 전환의 대표적인 핵심 기술 중 하나인 디지털 트윈에 대한 향후 전략적 연구방향 설정 및 도전 과제 발굴 등의 시사점을 도출하고자 한다.

1. 서론

미·중 무역갈등에서 촉발된 기술패권 전쟁, 러시아·우크라이나 전쟁 등 지정학적 긴장과 기후변화 문제의 심각성 강화, COVID-19 팬데믹 등 다양하고 복합적인 글로벌 환경은 전 산업에 전례없는 도전 과제를 던져주고 있다. 이와 같은 환경 변화로 인해 리쇼어링 등 국제 공급망의 불안정성이 커지고 원자재 가격 급등과 에너지 문제 등이 발생하면서 특히 제조 산업에 큰 타격을 주면서 전 세계는 경제적 위기에 직면하고 있다.

이러한 문제의 해결책을 찾기 위해 주요국들은 이미 첨단 기술 분야에서 경쟁을 시작하면서 사회 전반의 디지털 전환(digital transformation, DX)이라는 목표로 혁신을 도모하고 있다. 그 중 디지털 트윈(digital twin) 기술은 디지털 전환의 핵심 기술로 각광받고 있으며, 제조 분야에서는 이미 도입의 효과 및 성공 사례가 다수 등장하고 있다. 이에 따라, 제조 분야뿐만 아니라 의료, 건설, 에너지, 환경 등 다양한 응용 분야로 디지털 트윈 기술을 적용하고 고도화하면서 기술시장이 급격히 확대되고 있는 추세이다.

주요국들은 디지털 트윈 등 제조혁신 기술을 선도하기

위해 공격적으로 관련 정책 및 투자를 확대하고 있으며 국내에서도 최근의 ‘신 디지털 제조혁신 추진전략(2023.09)’ 등 노력을 이어나가고 있다. 하지만 국내 기술 수준은 최고기술을 보유하고 있는 미국에 비해 82.3%로 1.4년의 격차를 보이는 등(정보통신기획평가원, 2021) 디지털 트윈 기술의 경쟁력이 다소 취약한 상황으로 장기적 관점의 기술 내재화 전략이 필요한 시점이다.

이에 본 연구는 토픽 모델링 기법을 활용하여 그 동안 수행되어 온 디지털 트윈 문헌을 중심으로 연구 동향을 체계적으로 파악하여, 중장기적 관점의 디지털 트윈 기술 전략 수립 및 방향성을 제시하고자 한다.

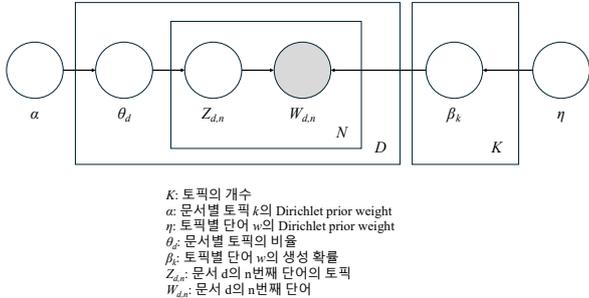
2. 관련 연구

2.1 토픽모델링

토픽 모델링(topic modeling)은 대량의 문서 집합에서 각 문서별 단어들의 분포를 확률적으로 구조화하여 문서별 토픽을 추론해내는 방법(Griffiths et al., 2004), 자연어 처리 분야, 비정형 텍스트 데이터 분석 분야에서 기술이나 연구 동향을 파악하는데 중요한 역할을 하는 분석

방법이다.

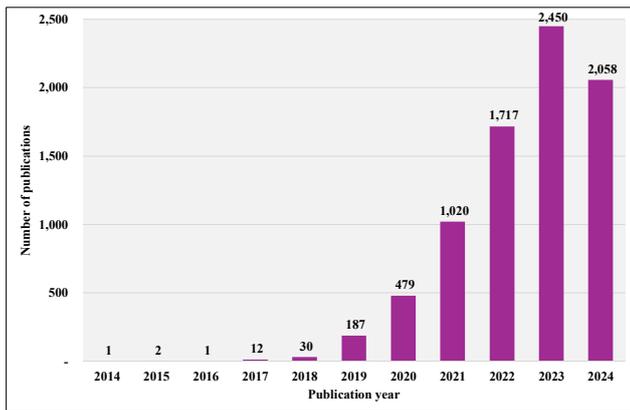
토픽 방법론 중 가장 대표적인 방법론은 LDA (Latent Dirichlet allocation) 모델로 서로 다른 문서가 다른 주제를 가리킨다면 각 문서에서 사용되는 단어는 서로 다른 분포를 가진다는 전제에서 시작된 기법이다(그림 1).



(그림 1) LDA 모델 (Blei, 2012)

2.2 디지털 트윈

한 분야의 연구가 성숙기에 접어들게 되면 연구 동향을 파악하여 연구가 나아가고 있는 방향이나 주제 등 특이점에 대한 분석이 가능해진다(서창교, 2020). 이러한 동향 파악을 통해 새로운 연구 방향성을 포착하거나 주요 주제 기반의 또 다른 논의를 이어나갈 수 있게 된다. 디지털 트윈을 대상으로 한 연구 논문을 통해 서지정보학 기반의 통계학적 특성 및 동향을 파악하고자 한 연구들이 다수 진행되었다. 연구 주제 및 출판물의 통계적 특성과 시계열적 변화 방향, 응용 분야를 분석하고 시사점을 제시한 연구들이 주를 이루고 있다(Gao et al., 2024; Ante, 2021).



(그림 2) 연도별 디지털 트윈 관련 논문 게재 수

하지만 디지털 트윈 관련 연구 문헌은 다른 주제에 비해 비교적 최신의 기술로 논문이 폭발적으로 증가하고 글

로벌 빅테크 기업(Ansys, Siemens)이 상용제품을 출시하기 시작한 2021년 이래 올해로 불과 4년밖에 되지 않았다(그림 2). 따라서 지속적으로 최신의 문헌들을 포함하여 새로운 연구 결과를 도출해내야 의미있는 분석 결과 및 시사점이 제시될 수 있다.

3. 연구 방법론

본 연구에서는 최신의 디지털 트윈 관련 문헌을 포함하여 연구를 진행하고자 한다. 이를 위한 데이터 수집을 위해 Web of Science (WoS)에서 ‘디지털 트윈’을 키워드로 설정하고 검색식(query)를 구성하였다. 2014년부터 2024년까지 총 8,041편의 논문이 수집되었고 그 중 토픽 모델링 분석을 위한 초록이 누락된 논문을 제외하고 7,957편의 논문을 최종 분석에 활용하였다. 해당 데이터를 대상으로 토픽 모델링 방법론 중 LDA 기법을 활용하여 연구 동향을 파악하고자 한다.

참고 문헌

[1] 서창교, “토픽모델링과 소셜 네트워크 분석을 이용한 공급사슬관리 연구 탐색”, 한국SCM학회지, 제20권 제1호, 2020, pp.36-51.

[2] 정보통신기획평가원, 디지털 트윈의 기술적 정의와 세부적 발전 5단계(level) 모델, 2021.

[3] Ante, L., “Digital twin technology for smart manufacturing and industry 4.0: A bibliometric analysis of the intellectual structure of the research discourse”, Manufacturing Letters, Vol.27, 2021, pp.96-102.

[4] Blei, D. M., “Probabilistic topic models,” Communications of the ACM, Vol.55, No.4, 2012, pp.77-84.

[5] Gao, L., Li, S., & Guo, H., “Visualization and Analysis of Research Areas and Hot Topics in the Field of Digital Twins”, IEEE Access, Vol.12, 2024, pp.26897-26904.

[6] Griffiths, T. L., & Steyvers, M., “Finding scientific topics,” Proceedings of the National academy of Sciences (PNAS), Vol.101, 2004, pp.5228-5235.

동적 벡터 색인을 이용한 지능형 학술검색

선충녕, 이태석, 한상준, 유수현
한국과학기술정보연구원 융합서비스센터
wilowisp@kisti.re.kr, tsi@kisti.re.kr, sjhan@kisti.re.kr, yoosu@kisti.re.kr

Intelligent Academic Search Using Dynamic Vector Indexing

Choong-Nyoung Seon, Taeseok Lee, Sangjun Han, Suhyeon Yoo
KISTI Convergence Service Center

요약

학술 문헌의 양이 급증함에 따라 효율적인 학술 정보 검색에 대한 요구가 증가하고 있다. 키워드 기반 검색은 빠르고 간편하지만, 사용자 의도를 반영한 결과 제공이 어려운 한계가 있다. 최근에는 검색에 의도를 반영하기 위해 벡터를 이용하여 유사한 문서를 검색하고 거대 언어모델을 활용하여 응답을 제공하는 다양한 서비스가 등장하고 있다. 하지만 학술 문헌의 규모가 크고, 새로운 문헌이 꾸준히 추가 되기 때문에 문헌을 벡터로 색인하는 작업은 서비스에 부담이 된다. 이러한 한계를 극복하기 위해 본 논문에서는 검색 증강 생성을 실시간 색인과 결합하여 학술 정보 시스템에서 검색을 지능화하는 새로운 접근 방식을 제안한다. 제안된 방법은 벡터 기반 의미검색의 문맥 이해를 이용하면서도, 대규모 문헌에 대한 벡터 색인 부담을 낮춰 빠른 서비스 개발을 가능하게 한다.

1. 서론

최근 학술 연구 환경은 정보의 폭발적 증가와 함께 빠르게 변화하고 있다. 방대한 양의 문헌 속에서 연구자가 필요로 하는 정확한 정보를 신속하게 찾아내기 위해 검색의 중요성은 커지고 있다. 기존의 키워드 기반 검색은 빠르고 간단하게 검색할 수 있는 장점이 있지만, 검색 질의 의도를 명확하게 파악하지 못하기 때문에 부정확한 검색 결과를 제공할 수 있다.

전통적인 검색의 한계를 극복하고자 인공지능의 발전과 함께 거대언어모델(Large Language Models, LLMs)을 활용하고 검색을 이용해 답변을 생성하는 검색증강생성(Retrieval-Augmented Generation, RAG) 기법을 통한 서비스가 등장하고 있다. RAG에서는 정확한 정보를 LLM에 제공하는 것이 중요하므로, 기존의 키워드 검색이 아닌 의미기반 벡터 검색을 이용하여 문맥으로 제공하는 것이 일반적이다. RAG와 LLM을 이용한 학술 정보 검색은 자연어 처리 능력을 바탕으로 연구자들의 질의를 더 깊이 이해하고, 문헌 내에서 의미 있는 정보를 추출할 가능성을 제공한다. 그러나 벡터 기반 검색은 대규모의 학술 문헌을 벡터화하고 색인하는 과정에서 막대한 계산 자원을 소모하며, 새롭게 추가되는 문헌들에 대한 실시간 색인도 부담이 되기 때문에 학술 정보 검색 시스템의 구축과 운영에 있어 주요한 걸림돌이 되고 있다. 따라서 최근 기법들을 사용하는데 벡터 색인의 문제를 해결하는 것이 필요하다.

2. 관련 연구

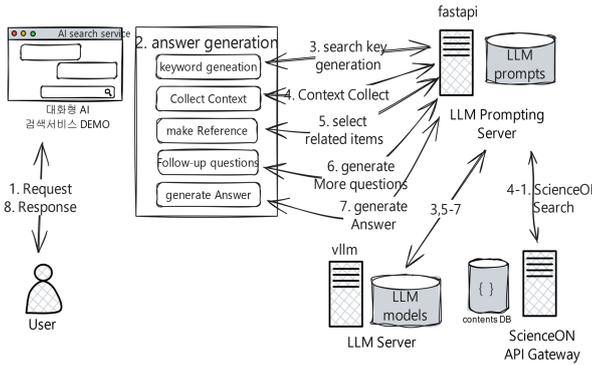
학술 정보 검색 분야에서는 키워드 기반 검색의 한계를

극복하고 보다 정확하고 의미 있는 결과를 제공하기 위한 다양한 연구가 진행됐다. Karpukhin은 Dense Passage Retrieval (DPR) 기법을 제안하여 대규모 문헌에서 효과적으로 관련 문서를 검색하는 방법을 소개했다[1]. DPR은 질의와 문서를 고밀도 벡터 공간에 매핑하고, 이들 간의 유사도를 계산하여 가장 관련성 높은 문서를 검색하여 전통적인 키워드 기반 검색보다 더 정확한 결과를 제공할 수 있으며, 특히 복잡한 질의나 문맥 의존적인 정보 검색에 효과적이다. Lewis는 Retrieval-Augmented Generation (RAG) 모델을 제안하여 지식 집약적인 자연어 처리 작업의 성능을 크게 향상시켰다[2]. 이 연구는 대규모 언어모델의 한계를 극복하고, 외부 정보를 효과적으로 활용하는 방법을 제시했다는 점에서 의의가 있다.

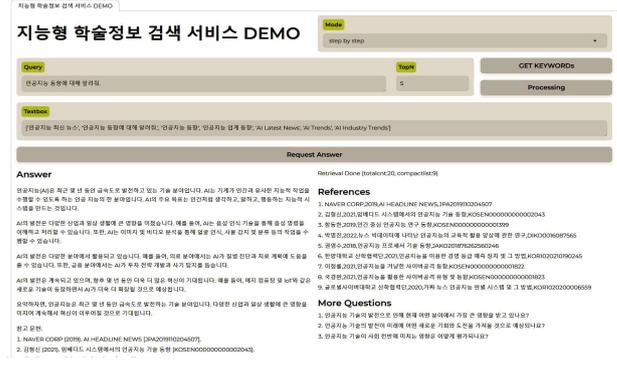
기존의 의미기반 벡터 검색 시스템들은 대규모 문헌을 벡터화하고 색인하는 과정에서 막대한 계산 자원을 필요로 한다는 한계가 있다. 본 연구에서 제안하는 방법은 전통적인 검색 엔진과 의미기반 검색을 결합하여 이러한 한계를 극복하고자 한다. 기존 검색 엔진의 빠른 처리 속도와 의미기반 검색의 정확성을 동시에 활용함으로써, 대규모 벡터 색인 과정을 최소화하면서도 높은 검색 정확도를 유지할 수 있는 접근 방식을 제안한다.

3. 키워드 검색 기반 지능형 학술정보 검색서비스

그림 1은 동적 색인을 적용하기 위한 지능형 학술정보 서비스의 구성 및 흐름을 보여준다. 사용자가 검색 질의를 입력하면, 응답 생성을 위한 클라이언트는 LLM prompt 중계를 위한 API server와 통신을 통해 기능을 수행한다.



(그림 1) 지능형 학술정보 검색서비스 시스템 구성



(그림 2) 지능형 학술정보 검색서비스 개념증명(Poof of Concept)

각 단계는 다음과 같은 순서로 수행된다.:

- (1) 키워드 검색을 위한 검색 키워드 확장 생성
- (2) RAG를 위한 검색 결과 수집
- (3) 관련 문서 선정
- (4) 추가 질의 생성
- (5) 응답 생성

검색 결과 수집을 위해 키워드 기반의 학술정보 검색을 제공하는 ScienceON API Gateway를 사용하고 있으며 이를 이용해 논문, 특허, 보고서, 동향의 자료를 검색한다. 관련 문서 선정은 모아진 키워드 검색 결과에 대해 의미적으로 관련 있는 문헌들을 정리하는 단계이며, 응답 생성은 문맥과 관련 문서로 사용자에게 생성 기반으로 응답을 제공하는 단계이다. 키워드 생성, 응답생성, 추가 질의 생성에 이용한 LLM은 KONI를 이용하였다[3].

4. 동적 색인을 이용한 의미기반 재 순위화

의미기반 검색을 위한 색인의 한계를 극복하기 위해 키워드 검색을 통한 후보 선택과 동적 색인을 통한 의미기반 재 순위화라는 두 단계로 후보 문서를 선별한다.

사용자의 요청은 키워드 검색에 부적절한 경우가 많아서 사용자 질의를 검색 형태로 변환하는 과정이 필요하다. LLM을 통해 사용자 질의와 관계된 검색 키워드 목록으로 확장하여 ScienceON API Gateway으로 검색 결과를 수집한다. 수집된 결과는 chunking와 embedding 과정을 통해 벡터화를 수행하게 된다. 이때 embedding model은 다국어에서 좋은 성능을 보여주고 있는 bge-m3[4] 모델을 사용하였으며 chunk의 벡터 결과를 저장하기 위해 Chroma DB를 사용하였다.

논문, 특허, 보고서, 동향에서 확장 키워드로 검색된 결과 중 상위 n개를 대상으로 응답 후보로 선정하고, 이들을 대상으로 동적 색인을 한 후 검색은 최대 한계 관련성 (Maximum Marginal Relevance, MMR)을 사용하였다. local LLM들의 context 크기를 고려하여 8k 수준으로 나올 수 있도록 최대 10개의 chunk를 목록화하도록 하였으며, 이 중 같은 문서의 chunk들은 취합하여 최종적으로 연관 문서들을 순위화하였다.

5. 결론

LLM을 이용하여 정보에 접근하는 다양한 서비스가 등장하고 있다. 특히 RAG와의 결합을 통해 환각 문제를 극복하고, 활용 가능한 서비스로의 가능성을 보여주고 있다. 하지만 대규모의 문헌을 대상으로 할 때 의미검색을 위한 벡터 색인 문제로 한계가 있었다. 이에 본 논문에서는 대규모 학술 문헌을 효과적으로 검색하기 위한 새로운 접근 방식을 제안하였다. 키워드 기반 검색과 동적 색인을 결합하여, 기존의 검색 방법이 가진 한계를 극복하고자 하였다. 제안된 방법은 검색 질의의 의도를 보다 정확하게 반영하면서도, 벡터 색인의 부담을 줄여 실시간 학술 정보 검색 서비스를 구축하는 데 있어 효율성을 향상시켰다. 특히, 의미기반 검색의 성능을 유지하면서도 대규모 문헌을 다루는 데 필요한 연산 자원을 최적화함으로써, 실용적이고 확장 가능한 검색 시스템을 구현할 수 있음을 확인하였다.

Acknowledgement

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다. (과제번호:K24L1M3C4)

참고문헌

- [1] Karpukhin, Vladimir, et al. "Dense Passage Retrieval for Open-Domain Question Answering." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Nov. 2020, pp. 6769-6781. doi:10.18653/v1/2020.emnlp-main.550.
- [2] Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 9459-9474.
- [3] KISTI. KISTI-KONI/KONI-Llama3-8B-Instruct-20240729. 2024, Hugging Face, <https://huggingface.co/KISTI-KONI/KONI-Llama3-8B-Instruct-20240729>.
- [4] Chen, Jing, et al. "BGE M3-Embedding: Multi-Lingual, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation." arXiv preprint, arXiv:2402.03216, 2024.

YOLOv8-pose 모델의 P6 구조 및 SimAM 모듈 기반 성능 개선

조상호*, 조인휘
한양대학교*, 한양대학교
e-mail :xianghao1234@naver.com

Performance Improvement of YOLOv8-pose Model Based on P6 Structure and SimAM Module

Xianghao Zhao*, Inwhee Joe
Hanyang University*, Hanyang University

요약

본 논문은 P6 구조 도입과 SimAM 모듈 통합을 통해 YOLOv8-pose 모델의 성능을 개선한 방법을 제안하였다. 또한, 피라미드 구조 확장을 통해 다양한 크기의 객체를 효과적으로 처리할 수 있는 능력을 강화하였다. 실험 결과, 제안된 모델은 기존 YOLOv8-pose 모델보다 mAP 지표에서 더 높은 성능을 보였으며, 특히 복잡한 배경과 다양한 자세 변화에서도 우수한 정확도를 유지하였다.

1. 서론

인간 자세 추정은 컴퓨터 비전 분야에서 중요한 연구 주제 중 하나로, 다양한 응용 분야에서 활용될 수 있다. 예를 들어, 인간 행동 인식과 가상 현실, 증강 현실 등의 기술은 인간의 자세를 정확히 추정하는 데 크게 의존한다. 최근 몇 년 동안, 딥러닝의 발전은 이러한 자세 추정 문제의 정확도와 효율성을 크게 향상시켰다. 특히, YOLO 시리즈, 즉 You Only Look Once 모델은 실시간 성능과 높은 탐지 능력 덕분에 컴퓨터 비전에서 매우 중요한 도구로 자리 잡았다 [1]. 그 중에서도 YOLOv8-pose 모델은 자세 추정 작업을 위해 최적화된 버전으로, 기존 YOLO 모델의 장점을 유지하면서도 보다 정교한 자세 추정을 가능하게 한다. YOLOv8-pose 모델은 자세 추정 작업에서 빠른 처리 속도와 높은 정확도를 제공하지만, 여전히 복잡한 배경과 다양한 자세 변화에서 성능 향상을 위한 개선의 여지가 있다.

이러한 문제를 해결하기 위해 본 논문에서는 세 가지 주요 개선 사항을 제안한다. 첫째, YOLOv8-pose 모델의 성능을 개선하기 위해 P6 구조를 도입하였다. 둘째, SimAM, 즉 유사도 인식 주의 모듈을 통합하여 모델의 주의 메커니즘을 강화하였다 [2]. 셋째, 피라미드 구조를 확장하여 다중 스케일의 특성 융합을 최적화하였다 [3]. 이 연구는 OpenThermalPose 데이터셋의 적외선 이미지와 COCO 2017 데이터셋의 일부 가시광선 이미지를 결합하여, 다양한 조명 조건과 배경에서 YOLOv8-pose 모델의 일반화 능력을 평가하였다 [4][5]. 이러한 개선 사항들은 YOLOv8-pose 모델의 기존 한계를 극복하고, 다양한 응용 시나리오에서 더 높은 정확도와 강건성을 보장하기 위해 설계되었다.

2. 본론

본 연구에서는 YOLOv8-pose 모델의 성능을 향상시키기 위해 P6 구조를 도입하고, SimAM 모듈을 통합하였으며, 피라미드 구조를 확장하였다. 이러한 개선을 통해 모델의 다중 스케일 처리 능력과 주의 메커니즘이 강화되었으며, 다양한 크기의 객체와 복잡한 배경에서도 높은 정확도를 유지할 수 있도록 설계되었다.

2.1.P6 구조 도입:

P6 구조를 도입함으로써 기존 YOLOv8-pose 모델의 한계를 극복하고, 고해상도의 작은 객체와 세밀한 정보를 더 잘 탐지할 수 있도록 하였다. P6 구조는 기존의 P5 구조에서 확장된 계층으로, 더 높은 해상도의 특성을 추출하여 모델이 더 작은 객체를 인식하는 데 필요한 세부 정보를 더욱 정확하게 학습할 수 있게 한다. 이는 특히 복잡한 장면에서 작은 객체의 정확한 인식을 가능하게 하여, 자율주행, 보안 감시 등과 같은 실시간 응용 분야에서 중요한 역할을 한다.

2.2.SimAM 모듈 통합:

SimAM, 즉 유사도 인식 주의 모듈은 주의 메커니즘을 강화하여 중요한 특성에 대한 집중도를 높인다. 이 모듈은 모델이 입력 데이터의 각 특성 간의 유사성을 계산하여, 중요한 정보를 강조하고, 덜 중요한 정보는 억제하도록 설계되었다. 이를 통해 모델은 복잡한 배경에서도 높은 정확도를 유지할 수 있으며, 다양한 환경에서 더욱 안정적인 성능을 발휘할 수 있다 [2]. 특히, SimAM 모듈의 통합은 YOLOv8-pose 모델이 복잡한 환경에서 객체를 더 잘 구별하고, 혼잡한 장면에서의 객체 탐지 성능을 향상시키는 데 기여하였다.

2.3.피라미드 구조 확장:

피라미드 구조 확장은 다중 스케일에서 특성을 융합하여 다양한 크기의 객체를 동시에 처리하는 데 중요한 역할을 한다. 기존 모델의 피라미드 구조는 특정 크기의 객체에 대한 특성 추출에 제한적이었으나, 이를 확장함으로써 모델은 여러 스케일에서의 특성을 효과적으로 통합할 수 있게 되었다. 이를 통해 모델은 작은 객체와 큰 객

체를 모두 효과적으로 탐지할 수 있으며, 복잡한 장면에서도 우수한 성능을 유지할 수 있다 [3]. 이러한 구조적 개선은 특히 비정형 데이터가 많은 응용 환경, 예를 들어, 도심 환경의 자율주행 차량 탐지나 혼잡한 공공 장소의 보안 감시에서 그 효과를 발휘한다.

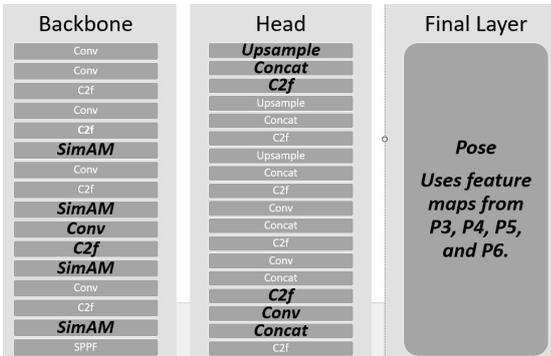


그림 1. YOLOv8-Pose 모델에서 P6 구조와 SimAM 모듈 통합을 통한 개선된 모델 구조

모델	POSE-Precision	POSE-Recall	POSE-mAP50	POSE-mAP50-95
기본 YOLOv8s-pose	0.891	0.749	0.791	0.526
개선된 YOLOv8s-pose	0.910	0.803	0.846	0.566

표 1. 기존 YOLOv8s-pose 모델과 개선된 YOLOv8s-pose 모델의 성능 비교

모델	POSE-Precision	POSE-Recall	POSE-mAP50	POSE-mAP50-95
기본 YOLOv8n-pose	0.900	0.739	0.775	0.484
개선된 YOLOv8n-pose	0.895	0.750	0.801	0.500

표 2. 기존 YOLOv8n-pose 모델과 개선된 YOLOv8n-pose 모델의 성능 비교

3. 실험 방법

본 연구에서는 YOLOv8-pose 모델의 성능을 평가하기 위해 OpenThermalPose 데이터셋과 COCO 데이터셋의 일부를 결합한 혼합 데이터셋을 사용하였다. [4][5]이 데이터셋은 적외선 이미지와 가시광선 이미지로 구성되어 있으며, 다양한 조명 조건과 복잡한 배경에서 모델의 자세 추정 능력을 평가하기 위해 활용되었다. 모든 모델은 동일한 조건에서 학습되었으며, 성능 비교는 기존 YOLOv8-pose 모델과 개선된 모델 간의 다양한 성능 지표를 기준으로 이루어졌다.

4. 결론

본 연구에서는 YOLOv8-pose 모델의 성능을 향상시키기 위해 P6 구조, SimAM 모듈, 및 피라미드 구조의 세 가지 주요 개선 사항을 제안하고 평가하였다. 실험 결과, 이러한 개선 사항들은 다양한 응용 시나리오에서 모델의 정확도와 강건성을 높이는 데 기여하였다. 특히, 개선된 모델은 복잡한 배경과 다양한 자세 변화에서도 우수한 성능을 유지하며, 이는 실시간 응용 분야에서 매우 중요한 요소로 작용할 수 있다. 향후 연구에서는 더 다양한 데이터 세트와 추가적인 개선 모듈을 통합하여 모델의 성능을 더욱 강화할 수 있을 것이다.

참고 문헌

[1] Terven J, Córdova-Esparza D M, Romero-González J A. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas[J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680-1716.

[2] Yang L, Zhang R Y, Li L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks[C]//International conference on machine learning. PMLR, 2021: 11863-11874.

[3] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.

[4] Kuzdeuov A, Taratynova D, Tleuliyev A, et al. OpenThermalPose: An Open-Source Annotated Thermal Human Pose Dataset and Initial YOLOv8-Pose Baselines[J]. Authorea Preprints, 2024.

[5] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context[C]//Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. Springer International Publishing, 2014: 740-755.

인공지능 개발 지원을 위한 통합 플랫폼 운영 전략

황명권^{1,2}, 최장원¹, 최선희¹
¹한국과학기술정보연구원 (KISTI)
²과학기술연합대학원대학교 (UST)
e-mail : {mgh, jwchoi, sunny.choi}@kisti.re.kr

A strategy for integrated platform to support artificial intelligence development

Myunggwon Hwang^{1,2}, Jangwon Choi¹, Seonhui Choe¹
¹Korea Institute of Science and Technology Information (KISTI)
²University of Science and Technology (UST)

요 약

KISTI는 국내의 연구 지원을 위해 슈퍼컴퓨팅 자원을 운영하고 있다. 최근 인공지능 개발에 대한 수요가 늘어감에 따라 데이터, 모델, GPU의 통합된 형태의 플랫폼 제공이 필요한 시점이다. 본 연구에서는 KISTI가 보유하고 있는 자원과 서비스를 연계하고 국내 AI 개발 지원을 위한 통합 플랫폼에 대해 제안한다.

1. 서 론

인공지능이 모든 영역에서 활용됨에 따라 GPU 클라우드의 사용 규모 또한 지속적으로 증가하고 있는 추세이다. 2024년 8월에 발행된 Fortune Business Insights의 보고서에 따르면, 클라우드 형태로 제공되는 GPUaaS(GPU as a Service)의 북미 시장 규모가 2023년 32억 3천만 달러로 평가되었다.[1] 또한 동일한 보고서에서 2024년 43억 1천만 달러로 예측하였으며, 2032년에는 498억 4천만 달러로 성장할 것으로 보고하고 있다. 국내의 경우에도 향후 5년간 연평균 성장률 9.9%를 기록할 것이라는 예측이다.[2] 지속적으로 늘어나는 수요에 대응하기 위해 KISTI는 슈퍼컴퓨터 6호기 도입을 위해 노력하고 있다. 또한 KISTI는 과학기술과 관련된 데이터, AI 모델 등 인공지능 개발에 필요한 필수 영역에서의 자원과 서비스를 제공하고 있다. 본 논문에서는 국내의 AI 개발 수요에 대응하기 위해 KISTI가 보유하고 있는 자원과 서비스들을 통합한 형태로 제공할 수 있는 플랫폼 운영 전략을 기술한다.

2. KISTI 보유 인프라 및 서비스

2.1 데이터

2024년 1월 1일을 기준으로 KISTI가 보유하고 있는 데이터에 대한 정보는 [표 1]과 같다. 데이터는 AI의 개발을 위한 필수 자원이기 때문에, 이를 보유하고 있다는 것은 과학기술 분야 전문가 AI를 만드는데 있어서 강점으로 작용한다. [표 1]의 데이터들은 ScienceOn¹)에서 통합적으로 검색과 OpenAPI를 통해 제공하고 있다.

(표 1) KISTI 보유 데이터

종류	건수
국가 R&D 보고서 원문	257,107건
OA논문	42,874,441편 (국내 1,604,826편, 국외 41,269,615편)
국내외 연구데이터 수집 및 연계	국내 40,111 건 국외 1,887,541 건

2.2 컴퓨팅 인프라

현재 보유하고 있는 슈퍼컴퓨터는 5호기에 해당한다. 그 성능은 25.71PFlops로 2024년 6월에 발표된 TOP-500을 기준으로 75위에 해당한다. 현재 최신 수요를 반영한 6호기 도입을 위한 계획이 있으며, 600PFlops 이상의 성능을 목표로 한다.

2.3 인공지능 모델

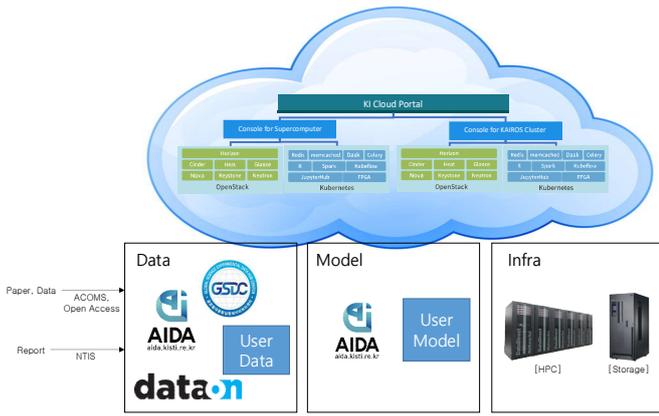
KISTI의 서비스 중 하나인 AIDA²)에서는 인공지능 학습용 데이터와 모델을 제공하고 있다. 이 서비스는 KISTI에서 구축한 AI 데이터를 자유롭게 접근하여 처리/분석하고 연구 결과를 도출할 수 있도록 지원한다. 현재 보유하고 있는 데이터 셋은 텍스트, 이미지 등 22개이며, 모델은 KISTI 대표 과학기술 특화 언어모델인 KONI³)를 비롯하여 23개를 제공하고 있다.

1) <https://scienceon.kisti.re.kr/main/mainForm.do>

2) <https://aida.kisti.re.kr/>

3) KONI-Llama3-8B-Instruct-20240729 :

<https://aida.kisti.re.kr/model/bd5c97dd-e331-4682-937c-4110597eb96f>



(그림 1) 통합 플랫폼 개요도

3. 인공지능 개발 지원을 위한 통합 플랫폼 운영 전략

3.1 통합 플랫폼 구성도

KISTI가 제공하는 다양한 서비스를 통해 인공지능 연구자의 AI 개발을 지원하기 위해서는, 개별적인 서비스 제공은 비효율적이라 할 수 있다. 현존하는 대중 서비스의 경우에도 데이터, 모델, 인프라를 별도로 제공하고 있으며, 이를 연계하는 부담은 개발자가 코드레벨에서 진행해야 하는 부담으로 남아있다. 이를 개선하기 위해, (그림 1)과 같이 데이터-모델-슈퍼컴퓨팅 인프라를 하나의 서비스처럼 제공하고, 사용자는 그 상위레벨에서 개발에만 집중할 수 있는 형태의 플랫폼을 제안한다.

3.2 모듈별 설명

가장 먼저 데이터를 다루는 영역이다. KISTI는 보고서, 논문 등 과학기술 정보를 다루는 데이터에 대한 서비스와 첨단 연구 장비에서 생산되는 대용량 데이터에 대한 공유/활용 플랫폼을 제공한다. 이들을 통합하여 사용자가 학습/분석을 위해 선택할 수 있도록 통합된 형태로 제공되어야 한다. 이를 위해 AIDA, DataOn, GSDC⁴⁾ 등이 사용자 레벨에서 하나처럼 보이고 선택할 수 있어야 한다. 사용자가 원하는 데이터가 없을 경우 직접 업로드 하고 메타데이터를 작성할 수 있으며, 이에 대한 공유 권한 설정이 가능해야 한다.

데이터가 선택되면, 그 이후에는 모델 선택부분이다. 모델은 AIDA 서비스를 통해 제공되고 있다. 현재는 사용자가 다운로드하고, 활용하기 위한 코드를 직접 작성해야 하지만, 이 또한 간단한 UI를 통해 쉽게 선택할 수 있어야 한다. 원하는 모델이 없다면 사용자 모델의 업로드와 외부 모델과 연계가 용이하게 제공되어야 한다.

현재 슈퍼컴퓨팅 인프라의 제공은 가장 활성화되어있는 영역이라 할 수 있다. 하지만 기존의 독립적인 운영 방식에서 데이터-모델-인프라를 하나의 UI에서 선택하고 사용

자의 부담을 최소화하며, KISTI 전체의 자원과 서비스를 이용할 수 있도록 통합된 플랫폼 형태로 제공되어야 한다.

이러한 통합된 플랫폼을 제공하기 위해 세부적으로는 데이터 수집 및 전처리, 데이터 저장, 버전관리, 데이터 가상화, MLflow, Kubeflow, Docker, Kubernetes, MLOps, GPU 가속, 서버리스 컴퓨팅, Low-Code/No-Code 등의 기술 연구 또한 병행해야 한다.

통합 플랫폼을 통해 AI를 개발하는 사용자는 하나의 UI에서 모든 자원을 접속할 수 있어야 하며, 동시에 모델 학습까지도 쉽게 진행될 수 있어야 한다.

4. 결론

본 논문에서는 국내 인공지능 개발 연구자를 위해 KISTI가 제공하는 다양한 자원과 서비스의 통합된 형태의 플랫폼에 대한 제언을 기술하였다. 기존의 방식은 데이터/모델의 수집과 제공, 슈퍼컴퓨팅 인프라의 구축과 제공이 개별적으로 제공되었지만, 슈퍼컴퓨팅 자원 6호기가 도입 예정인 현 시점과 인공지능 개발 수요의 급증한 증가에 맞물려서 최대 효율성/활용성을 확보할 수 있는 전략 중의 하나가 통합 플랫폼이라 할 수 있다. 사용자의 이용 관점에서 하나의 UI가 데이터-모델-인프라 활용, 그리고 코드작성과 학습까지 용이하게 가능하도록 해야 할 것이다.

감사의 글

이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 자체 사업으로 수행된 연구입니다.(과제번호: J24JR042-241)

참고 문헌

- [1] Hardware & Software IT Services / GPU as a Service Market - Fortune Business Insights : <https://www.fortunebusinessinsights.com/ko/gpu-as-a-service-market-107797>
- [2] <https://www.datanet.co.kr/news/articleView.html?idxno=193147>

4) GSDC (Global Science experimental Data hub Center)

멀티모달 프롬프트 학습에 대한 재고하기 비전 언어 모델

저우형^{1,2} 김상철^{*,2}

¹전북대학교 전자정보공학부

²전북대학교 지능형로봇연구소

hengz@jbnu.ac.kr, *sckim7777@gmail.com

Rethinking of Multi-Modal Prompt Learning for Vision-Language Models

Heng Zhou^{1,2}, SangCheol Kim²

¹Department of Electronics and Information Engineering, Jeonbuk National University

²Core Research Institute of Intelligent Robots, Jeonbuk National University

Abstract

This paper introduces multi-modal prompt learning, a novel technique that extends prompt-based approaches from natural language processing to vision-language models (VLMs). By leveraging task-specific prompts to guide model behavior, multi-modal prompt learning enables more flexible and efficient integration of visual and textual data, leading to improved performance on a range of tasks, including image captioning, visual question answering, and cross-modal retrieval. Unlike traditional methods that treat vision and language inputs separately, this approach fosters a unified learning process, allowing VLMs to better capture relationships between images and text. We highlight several key advantages of multi-modal prompt learning, such as enhanced task adaptability, reduced data requirements through prompt-driven fine-tuning, and more cohesive cross-modal representations. In addition, we propose three key research directions to further advance this field: (1) dynamic prompt generation for context-aware adaptation, (2) few-shot and zero-shot learning using multi-modal prompts to improve data efficiency, and (3) cross-modal prompt transfer for generalizing across tasks and modalities. These directions offer a promising roadmap for future innovations in multi-modal AI, expanding the applicability and robustness of VLMs across diverse domains.

Keywords: Computer Vision, Prompt Learning, Multi-modal Learning

1. Introduction

Recent advancements in artificial intelligence have led to the development of multi-modal models that can process and integrate information from different modalities, such as vision and language. Among these, Vision-Language Models (VLMs)

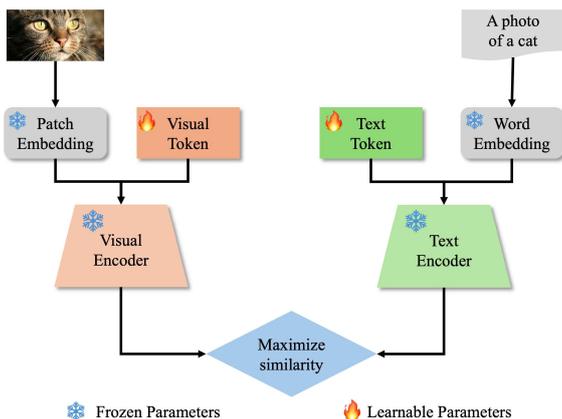
[1] [2] stand out as a powerful paradigm that combines the strengths of visual and textual data to perform a wide range of tasks, including image captioning [3], visual question answering [4], and cross-modal retrieval. One of the key methods driving the success of these models is multi-modal prompt learning, a technique inspired by natural language prompts used in large language models. In natural language processing (NLP), prompt-based learning involves providing models with input in a structured prompt format, guiding the model's behavior or prediction in response to a task. This concept has been extended to multi-modal domains, where prompts are designed to condition the model's understanding of both visual and textual inputs. The core idea behind multi-modal prompt learning is to craft a unified representation that allows the model to seamlessly process and

fuse information from different modalities, enabling more nuanced and flexible responses.

Multi-modal prompt learning offers several advantages. Firstly, it is able to increase task adaptability. By using task-specific prompts, VLMs can adapt to various downstream tasks without requiring large-scale retraining. For instance, a VLM might be prompted to generate a description of an image in one task, while in another, it could answer questions based on the same image. Secondly, it can improve data efficiency. Prompt-based approaches often require less data compared to traditional training methods because they leverage pre-trained models and focus on fine-tuning or conditioning through prompts. This allows for more efficient use of available datasets across multiple domains. Finally, it provides a unified learning for VLMs. Unlike earlier multi-modal systems that handled visual and textual inputs separately, prompt learning facilitates a more integrated learning approach. By fusing modalities at an early stage, VLMs achieve a deeper understanding of the relationships between images and text.

2. Related works

In this paper, we first provide a detailed description of multi-modal prompt learning, introducing its core concepts and computational processes. Following this, we discuss potential research topics within this area of study. Figure 1 illustrates the classic framework of multi-modal prompt learning, which includes two encoder layers, as well as visual embedding and text embedding layers. Notably, the parameters of these layers remain frozen during the model's fine-tuning process. The input to the model consists of an image-text pair, where the text follows the template "a photo of a category", with "category" representing the image's category, such as "cat" in this instance. Trainable visual tokens and text tokens are inserted into the visual and text embedding layers, respectively. These tokens, along with the original tokens, are passed through the two encoder layers for feature extraction. During training, backpropagation is performed by maximizing the similarity between the visual and textual features. In the testing phase, the textual features are used as weights to compute the logits of the two features, and classification is performed using softmax.



(Figure 1) The structure of multi-modal prompt learning

3. Future research

The first topic focuses on dynamic prompt generation for context-aware learning. A key challenge in multi-modal prompt learning is creating prompts that can adapt to varying tasks and inputs. Research in this area could explore techniques for generating prompts dynamically, rather than relying on pre-defined ones. This would involve the model analyzing the context of an image (e.g., objects, scenes, interactions) to generate appropriate prompts, enabling more accurate and meaningful responses for complex tasks such as scene understanding or video summarization. The second topic is few-shot and zero-shot learning using multi-modal prompts. While prompt learning has proven effective in fine-tuning pre-trained models with limited data, applying it to few-shot and zero-shot learning could extend its use to low-resource environments. Researchers could explore how prompt

engineering might allow VLMs to perform well in novel tasks or domains with minimal or no labeled data. This research would focus on developing generalized prompt architectures that can transfer knowledge across tasks and modalities without extensive retraining. The final topic is cross-modal prompt transfer and adaptation. While cross-modal transfer learning is gaining traction, little research has been done on cross-modal prompt transfer. This area would investigate how prompts learned in one modality (e.g., language tasks) can be adapted to another (e.g., visual tasks). Understanding how prompts generalize across modalities could lead to more efficient multi-modal models with less task-specific customization. Additionally, this research could explore how multi-modal prompts can address increasingly complex tasks involving deeper integration of vision and language.

Acknowledgment

This research was supported by the University-Focused Research Institute Support Program of the National Research Foundation of Korea (NRF-2019R1A6A1A09031717) and was funded by the Smart Farm Research and Development Project (No. 421023-04) which is part of financial resources of 2023 government (Ministry of Agriculture, Food and Rural Affairs, Ministry of Science and ICT, Rural Development Administration).

References

- [1] Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J. and Krueger, G., 2021, July. Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PMLR.
- [2] Jia, C., Yang, Y., Xia, Y., Chen, Y.T., Parekh, Z., Pham, H., Le, Q., Sung, Y.H., Li, Z. and Duerig, T., 2021, July. Scaling up visual and vision-language representation learning with noisy text supervision. In International conference on machine learning (pp. 4904-4916). PMLR.
- [3] Li, J., Li, D., Savarese, S. and Hoi, S., 2023, July. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In International conference on machine learning (pp. 19730-19742). PMLR.
- [4] Parelli, M., Delitzas, A., Hars, N., Vlassis, G., Anagnostidis, S., Bachmann, G. and Hofmann, T., 2023. Clip-guided vision-language pre-training for question answering in 3d scenes. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5607-5612).

중소기업 기업가 지향성이 비즈니스모델 혁신과 기업가적 성과에 미치는 영향: 조직문화 조절효과

곽영*, 유화선*, 정도범**

한국과학기술정보연구원 정책전략본부 정책연구센터*, 연구전략센터**

young.kwak@kisti.re.kr; hsyoun@kisti.re.kr; dbchung@kisti.re.kr

A study on the influence of SME entrepreneurial orientation on business model innovation and entrepreneurial performance

Young Kwak*, Hwasun You*, Dobum Chung**

Korea Institute of Science & Technology Information Div of Policy and Strategy

요약

본 연구는 빠르게 변화하는 경영환경 속에서 중소기업의 기업가 지향성, 비즈니스모델, 조직문화를 통해 기업가적 성과를 높이는 요인과 관계를 알아본다. 이를 통해 향후 중소기업이 시장변화에 적응하여 지속적인 성장과 경쟁력 확보를 유지할 수 있는 실질적인 경영전략을 제공한다.

1. 서론

빠르게 변화하는 사회 속에서 기업들은 글로벌화, 기술발전, 치열한 경쟁, 그리고 고객 욕구의 다양화 등 급격히 변화하는 환경에 직면해 있다[1]. 이러한 변화는 기업이 끊임없이 새로운 사업 기회를 찾아내고, 시장변화에 신속하게 대응할 수 있는 기업가 지향성을 필수적으로 갖추도록 요구하고 있다. 아울러, 미래의 경쟁 우위를 확보하기 위해 기존의 비즈니스모델을 재해석하고, 이를 바탕으로 새로운 비즈니스 개념 창출하는 능력이 요구되고 있다[2]. 이러한 경쟁 우위를 강화하기 위하여 지식을 창출하고 활용하는 조직의 폭넓은 활용은 기업 성과에 영향을 주는 근본요소로서 조직문화가 있다. 이러한 관점에서 기업가 지향성, 비즈니스모델 혁신, 그리고 이를 뒷받침하는 조직문화가 어떻게 기업가적 성과에 영향을 미치는지를 규명할 필요가 있다. 이를 통해 이론적인 연구와 실무에 기여하고자 하는 것이 본 연구의 목적이다.

2. 관련 연구

2.1 기업가 지향성

기업가 지향성은 기업의 열정이나 의지, 기업가의 능력, 산업에서의 역할 등의 다양한 의미로 정의되었다[3]. 또한, 제품과 시장의 혁신을 추구하고, 어느 정도 위험을 감수하며, 진취적으로 대응하여 경쟁자의 기선을 제압하려는 조직의 성향이라고 할 수 있다[4]. 즉, 기업가 지향성은 기업으로 하여금 새로운 사업의 기회를 끊임없이 모색하게 하고, 경쟁자보다 먼저 주어진 자원을 혁신적으로 활용하여 기업 목적을 효과적으로 달성하게 하는 활동을 의미한다.

2.2 비즈니스모델 혁신

비즈니스모델은 가치 창출과 가치 포착을 설명하기 위

해 기업이 어떻게 운영되는지 설명하는 방법입니다[5]. 비즈니스모델 혁신은 기존에 존재하지 않았거나 존재하지 않았던 새로운 비즈니스모델을 개발하여 새로운 가치를 획득하기 위해 혁신적인 비즈니스모델을 만드는 것을 말한다. 이러한 관점에서 비즈니스모델 혁신은 비즈니스모델의 점진적인 변화가 아닌 전체 비즈니스에 영향을 미치는 급진적 또는 파괴적 혁신을 말한다.

2.2 기업가적 성과

기업의 실적은 일반적으로 매출 성장률, 투자 수익률, 총 자산 성장률 등의 재무적 실적의 객관적인 지표를 통해 표현된다. 그러나 기업가의 실적은 자금 투자 및 새로운 설비 도입과 같이 수익 창출이 이루어지기까지의 시간적 지연으로 인해 재무제표에서 수치적 수치만으로 측정하기 어렵다[6]. 따라서 객관적인 평가보다는 주관적인 평가가 목적에 적합한 정보를 제공할 수 있고, 정보의 신뢰성이 상대적으로 높고, 정량적 평가보다는 정성적 평가가 현실적이라는 점을 감안하여 주관적 성과측정의 필요성이 강조된다.

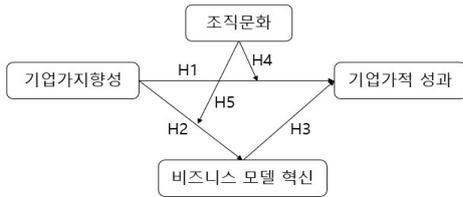
2.4 조직문화

조직문화는 특정 조직이 외부 환경에 적응하고 내부를 통합해 가면서 만들어 낸 특정한 양식으로 오랜 기간 조직구성원들에게 받아들여지고 학습되며, 조직구성원들이 세대를 거쳐 공유하고 이어온 조직의 문화, 의식, 환경 등의 집합을 말한다[7]. 조직문화는 지향점에 따라 관계, 혁신, 위계, 파업 지향으로 나눌 수 있으며 혁신적인 조직문화의 경우 새로운 아이디어, 프로세스, 신제품 등을 성공적으로 채택 또는 실행하는 능력을 키워 새로운 제품, 서비스를 만들어 낸다.

3. 연구방법

3.1 연구모형

본 연구는 선행연구를 기반으로 그림1과 같은 연구모형을 제시한다. 본 연구는 중소기업 기업가적 성과에 영향을 주는 요인으로 기업가 지향성과 비즈니스모델 혁신을 변수로 사용하였다. 변수 측정은 선행연구를 통해 기업가 지향성의 혁신성, 진취성, 위험감수성을 독립변수로 사용하고, 비즈니스모델 혁신을 매개변수, 조직문화를 조절변수로 사용하였다.



(그림 1) 연구모형

4. 실증분석

4.1 표본의 일반적 특징

본 연구는 (재)한국청년기업가정신재단(KoEF)에서 2022년에 수행한 기업가정신 실태조사 기업편의 8,559개 자료로 중소기업에 해당하는 2,518개의 표본을 사용하였다.

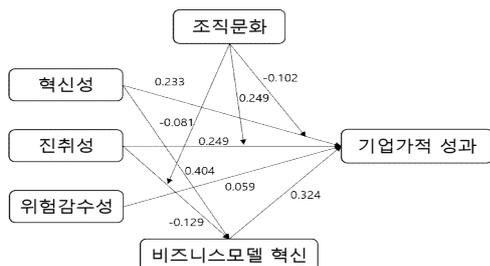
4.2 가설 검증

본 연구의 가설과 연구모형에서 제시한 변수 간의 인과관계를 분석하기 위해 SmartPLS 4.0을 사용하여 구조방정식(Structural Equation Modeling: SEM)을 적용하였다.

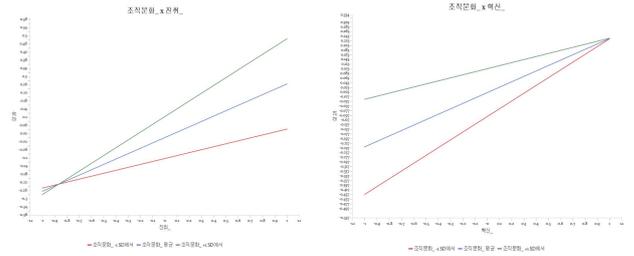
(표 6) 가설 검증

	경로계수	T Value	P Value	가설판정
위험→ 성과	0.067	3.452	0.001**	채택
위험→비즈	-0.035	1.527	0.127	기각
진취→성과	0.270	13.302	0.000***	채택
진취→비즈	0.238	9.440	0.000***	채택
혁신→성과	0.233	10.360	0.000***	채택
혁신→비즈	0.406	16.813	0.000***	채택
비즈→성과	0.331	16.577	0.000***	채택
문화×위험→성과	0.005	0.225	0.822	기각
문화×위험→비즈	0.007	0.257	0.797	기각
문화×진취→성과	0.109	5.214	0.000***	채택
문화×진취→비즈	-0.060	2.201	0.028**	채택
문화×혁신→성과	-0.066	2.954	0.003**	채택
문화×혁신→비즈	0.006	0.221	0.825	기각

***p < 0.001, **p < 0.05, *p < 0.01 수준에서 유의함



(그림 2) 경로계수와 가설검정



(그림 3) 조절효과 분석

4. 결론

5.1 연구결과 및 시사점

본 연구는 많은 중소기업들이 직면한 혁신과 기업가적 성과 달성에 필요한 기업가 지향성 요인을 파악하여 빠르게 변화하는 기업환경에 대비할 수 있도록 하였다. 또한, 높은 수준의 조직문화를 통해 비즈니스모델 혁신과 기업가적 성과를 이룰 수 있다는 결과를 확인할 수 있었다. 이러한 결과를 바탕으로 향후 중소기업이 지속적인 성장과 경쟁력 확보를 위한 실무적 시사점을 제안한다.

5.2 연구 한계점

향후 중소기업과 비슷한 환경의 사업초기 기업이나 벤처기업의 내외부 요인과 관련한 가설의 확장 연구가 필요하다고 판단된다. 또한, 성과에 대한 비재무적인 잠재 요인들을 명확히 함으로써 전략적 관리의 핵심 질문에 더 많이 기여할 수 있을 것이다.

참고문헌

[1] 이갑두(2012), “시장지향성과 기술지향성이 공급사슬 유연성과 성과에 미치는 영향: 경남지역기업을 중심으로”, 지역산업연구, 2012. 35(1). pp.81-82.

[2] IBM. 지속적 성장을 위한 비즈니스 모델 혁신전략. 2005. 서울: 한국경제신문.

[3] Schumpeter, J. A., "The theory of economic development : An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle" London : Oxford University Press, 1934.

[4] Miller, D., "The correlates of entrepreneurship in three types of firms," Management Science, Vol.29, No.7(1983), pp.770-790.

[5] Teece, D. J. (2010). Business models, business strategy and innovation. Long range planning, 43(2-3), 172-194.

[6] 박경미, 황재원. (2019). 경영전략과 최고경영자 특성이 기업성과에 미치는 영향. 한국산학기술학회 논문지, 20(1), 52-63.

[7] Schein, E. H., "Organizational Culture and Leadership, 2nd ed.", San Francisco, Jossey-Bass, 1996.

가제보 시뮬레이션 환경에서 자율형 4륜 관개 로봇의 견고한 경로 추적 및 제어

무함마드 이브라힘 자인 울 아비딘^{1,2} 김상철^{*,2}

¹전북대학교 전자정보공학부

²전북대학교 지능형로봇연구소

muhammad.ibrahim@jbnu.ac.kr, *sckim7777@gmail.com

Robust Path Tracking and Control for Autonomous Four-Wheel Irrigation Robots in a Gazebo Simulation Environment

Muhammad Ibrahim Zain Ul Abideen^{1,2}, SangCheol Kim²

¹Department of Electronics and Information Engineering, Jeonbuk National University

²Core Research Institute of Intelligent Robots, Jeonbuk National University

Abstract

This project focuses on the development and implementation of a path tracking system for a four-wheel irrigation robot using the Robot Operating System (ROS1) and the Gazebo simulation environment. The robot is designed to autonomously navigate through predefined paths in agricultural fields, ensuring efficient water distribution. Path tracking is achieved through the integration of sensor feedback, control algorithms, and kinematic models to accurately steer the robot along the desired trajectory. Various control strategies, such as Proportional-Integral-Derivative (PID) control and Model Predictive Control (MPC), are evaluated for their effectiveness in maintaining accurate path adherence. Simulation in Gazebo enables testing in diverse terrain conditions, facilitating performance analysis and optimization of the system prior to field deployment. The proposed system aims to enhance precision in irrigation tasks, contributing to water conservation and increased agricultural productivity.

Keywords: Path Tracking, Four-Wheel Irrigation Robot, ROS1, Gazebo Simulation, Autonomous Navigation.

1. Introduction

Recent advancements in robotics and automation have significantly impacted agricultural practices, leading to the development of autonomous systems capable of performing complex tasks with minimal human intervention [1]. One such innovation is the use of irrigation robots, which are designed to enhance water distribution efficiency in agricultural fields [2]. In particular, four-wheel irrigation robots have emerged as a practical solution, providing autonomous navigation through predefined paths while ensuring precise water application. This project focuses on the implementation of path tracking systems for such robots, leveraging the Robot Operating System (ROS1) and Gazebo simulation environment to achieve reliable autonomous movement.

Path tracking plays a crucial role in enabling robots to follow specific trajectories, ensuring that water is applied efficiently across the field. It involves the integration of sensor feedback, control algorithms, and kinematic models to maintain accurate path adherence [3]. By utilizing advanced control techniques like Proportional-Integral-Derivative (PID) and Model Predictive Control (MPC), the robot can adjust its

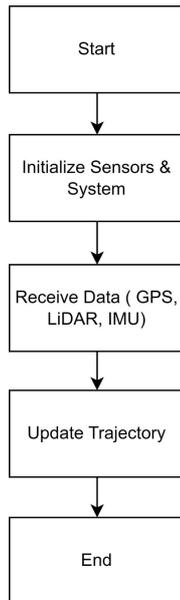
trajectory in response to environmental variations [4]. This not only enhances irrigation accuracy but also reduces resource consumption, contributing to increased agricultural productivity.

2. Related works

In this project, we first provide a comprehensive overview of path tracking techniques for autonomous robots, focusing on their application in agricultural environments. Previous research has explored various methods for autonomous navigation in both structured and unstructured environments. In particular, Proportional-Integral-Derivative (PID) control and Model Predictive Control (MPC) have been widely used for path tracking due to their simplicity and robustness. PID control has been effectively employed for its ability to correct deviations from a reference path by adjusting the robot's steering angle in real-time. On the other hand, MPC offers a more advanced approach by predicting future states of the robot based on its current position and velocity, allowing for smoother trajectory adjustments in response to obstacles or environmental changes.

Figure 1 illustrates a typical framework for path tracking in a four-wheel irrigation robot, where the robot's kinematic

model is combined with real-time sensor feedback to guide its movement along predefined paths. Various sensor types, such as GPS, LiDAR, and IMU, are used to continuously update the robot's position and orientation. Control algorithms, like PID and MPC, receive this data to adjust the robot's path, ensuring that it remains on course. During simulation in the Gazebo environment, the robot's behavior is tested under different terrain and weather conditions, allowing for the evaluation of the tracking system's performance. Testing in this controlled environment helps identify areas for optimization before the robot is deployed in real-world agricultural settings.



(Figure 1) The block diagram illustrates the key steps of the path tracking algorithm for a four-wheel irrigation robot, including sensor initialization, data reception, control application, and trajectory updates.

3. Future research

The first area of future work focuses on enhancing adaptive control strategies for varying field conditions. Current path tracking systems for agricultural robots, including the use of PID and MPC, operate effectively under predefined conditions. However, research could explore more dynamic control algorithms that adjust in real time to unpredictable field variables such as uneven terrain, varying soil moisture, and obstacles. By integrating machine learning techniques, future systems could allow the robot to learn optimal control actions based on real-time data, enabling smoother path adherence in complex environments.

The second area is the application of advanced sensor fusion techniques. While the current system relies on GPS, LiDAR, and IMU data, future work could investigate the fusion of additional sensors, such as cameras or thermal sensors, to improve accuracy and reliability. Sensor fusion could enhance the robot's ability to detect and respond to field anomalies,

such as crop health issues, leading to more efficient irrigation patterns.

Finally, future work could focus on the integration of cloud-based systems for large-scale deployment. By connecting multiple irrigation robots through a shared network, data from different robots could be aggregated to create a more comprehensive view of field conditions. This would allow for coordinated water distribution across large farms, optimizing resource use and improving overall agricultural efficiency. Additionally, cloud-based systems could facilitate remote monitoring and adjustments, making irrigation management more scalable and efficient.

Acknowledgment

This research was supported by the Regional Innovation Strategy (RIS) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2023RIS-008), the Agricultural Policy Support Technology Development Project funded by the Rural Development Administration's National Institute of Crop Science (RS-2023-00232338), and the University-Focused Research Institute Support Program funded by the National Research Foundation of Korea (NRF-2019R1A6A1A09031717).

References

- [1] Mahmud, M. S. A., Abidin, M. S. Z., Emmanuel, A. A., & Hasan, H. S. (2020). Robotics and automation in agriculture: present and future applications. *Applications of Modelling and Simulation*, 4, 130-140.
- [2] Rafi, R. H., Das, S., Ahmed, N., Hossain, I., & Reza, S. T. (2016, December). Design & implementation of a line following robot for irrigation based application. In *2016 19th International Conference on Computer and Information Technology (ICCIT)* (pp. 480-483).
- [3] Cai, Y., Zhan, Q., & Xi, X. (2012). Path tracking control of a spherical mobile robot. *Mechanism and Machine Theory*, 51, 58-73.
- [4] Nuhel, A. K., Al Amin, M., Paul, D., Bhatia, D., Paul, R., & Sazid, M. M. (2023, August). Model Predictive Control (MPC) and Proportional Integral Derivative Control (PID) for Autonomous Lane Keeping Maneuvers: A Comparative Study of Their Efficacy and Stability. In *International Conference on Cognitive Computing and Cyber Physical Systems* (pp. 107-121). Cham: Springer Nature Switzerland.

노인 생활 향상을 위한 상황인식 기반 스마트단말기 서비스 디자인 개선에 대한 연구

– 양로원에서 사용되는 스마트단말기를 중심으로 –

주자안, 경병표
공주대학교 대학원 게임디자인학과
zian1998zian@smail.kongju.ac.kr, kyungbp@kongju.ac.kr

Research on Improving Intelligent Terminal Service Design Based on Context Awareness to Enhance the Lives of Older Adults

–Focusing on Intelligent Terminals Used in Nursing Homes–

Zi-An Zhu, Byung-Pyo Kyung
Dept. of Game Design at Kongju National University

Abstract

This paper takes the smart terminal of the nursing home as the research object, and the elderly in the nursing home as the target user, aiming to improve the quality of life and happiness of the elderly in the nursing home based on the theory and method of user experience design. This paper reviews the relevant literature, sorts out the development of context-awareness theory and the process of context-awareness service, and analyzes and summarizes the current situation, characteristics and problems of institutional elderly care service. Combining the advantages and characteristics of context-awareness intervention in institutional elderly care service, the service context factors of institutional elderly care are divided into four categories: user, service, society, and device.

1. Introduction

Aging has become one of the social issues faced by the world today. Currently, both the number and proportion of the population aged 60 and over are increasing worldwide. This unprecedented growth rate is expected to accelerate further in the coming decades, especially in developing countries. Against the backdrop of rapidly advancing aging, the decrease in family size and the diminishing capacity of children to provide care and attention can no longer meet the older adult's needs, leading to an increasing number of families opting for more professional nursing home services.

This thesis primarily investigates the design strategies of intelligent terminals in nursing homes, with the main goals of enhancing user experience and promoting sus-tainable development. Through studying existing smart older adult care models and in-telligent terminals, it identifies and summarizes trends and limitations. Incorporating context-awareness theory, it analyzes the cognition and experience of the older adult, as well as the characteristics of the research subjects, to identify key contextual factors. By examining the impact of different contextual factors on user cognition and behavior, it proposes strategies for guiding healthy aging and the sustainable development of appli-cations and older adult care systems.

2. Context Awareness Summary

Context refers to the conditions of the surrounding environment during the occurrence of an event, including location, available resources, and the environment itself. It represents a describable overall state comprising numerous factors. Due to the diversity in scholars' research fields, context has been classified in various ways. In recent years, there has been a tendency to divide it based on a human-centered approach. Marreiros and others believe that context can be both the environment in which the subject currently exists and the historical environment related to the subject. This includes user context, physical context, computational context, and temporal context, recorded and stored in chronological order. The aim is to enable the system to process and predict user behavior by synthesizing information from past and present contexts.

Active context awareness pertains to the explicit information directly perceived by the device, such as location, time, temperature, and device information, whereas passive context awareness refers to the information that requires external events to trigger or user input, such as habits, education level, and preferences. The context-aware process is divided into three stages: perception, computation, and output. As users receive services, the feedback on contextual factors is

perceived and iteratively optimized, aiming to offer more accurate services to users.

3. Development status and problems of nursing homes

Institutional elder care, emerging in response to the societal needs for older adult care against the backdrop of an aging population, primarily relies on nursing homes as its main platform, boasting advantages such as professional medical resources and comprehensive nursing services. With changes in family structures and improvements in the level of services provided by these institutions, institutional elder care has been embraced by many families, becoming an irreplaceable component of the elder care service system.

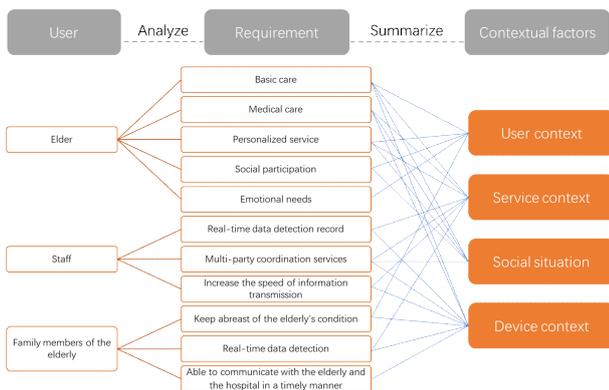
Through observational methods and interview surveys, this study investigates the service model of institutional elder care, identifying common characteristics and prevalent issues in the development of institutional elder care services.

(Table 1) Pain points

Pain points
Diverse Needs of Service Recipients and Unequal Resource Distribution.
The industrialization and specialization of institutional elderly care service.
Integration of Medical Care and Elder Care Services.
Application of Mobile Internet of Things (IoT) and Other Smart Information Technologies.

4. Context classification

Building on the development of context-awareness theory and integrating the advantages and characteristics of smart terminals facilitated by context awareness, a classification method for contextual factors suitable for smart terminals is summarized, along with an analysis of the connotation and relationships of each type of contextual factor.



(Figure) User needs and contextual factors

5. Smart Terminal Design Strategy Based on Context-Awareness Theory

The survey found that while the basic functional modules and service systems of nursing homes have gradually

developed and matured with technological support, the user experience of nursing home services still needs further enhancement. Thus, combining the user, service, social, and device contextual factors analyzed in this paper, corresponding smart terminal design strategies are proposed, and a smart elder care service system framework is constructed.

(Table 2) Smart Terminal Design Strategy Based on Context-Awareness Theory

User Context	Integrate multiple situations to provide personalized experience Improve user information transmission and collaboration efficiency
Service Context	Multi-user-Computer Status Identification Collaborative Service Model Improve user identification service quality
Social Context	Older adult care institutions intervene to form an auxiliary situation Increase social participation and make systems sustainable
Device Context	Multi-sensory interaction improves the perception of the older adult Provide intelligent machine fusion environment adaptive services

6. Conclusions

Nursing homes, as a critical pillar of the elder care system, play a vital role in mitigating the challenges posed by societal aging. Building on these factors and the user needs identified, design strategies tailored to the user, service, social, and device contexts are proposed. These strategies aim to increase the specificity and fit of elder care services, thereby improving the user experience within the service system. The theoretical framework of context-aware services offers new directions for the sustainable innovation and development of nursing homes.

References

[1]DESA U N. World Population Ageing 2020: Highlights (ST/ESA/SER. A/451). UN DESA: New York, NY, USA, 2020.

[2]Schmidt A. Context-aware computing: context-awareness, context-aware user interfaces, and implicit interaction. The Encyclopedia of Human-Computer Interaction, 2nd Ed., 2013.

[3]Marreiros G, Santos R, Ramos C, et al. Context-aware emotion-based model for group decision making. IEEE Intelligent Systems, 2010, 25(2): 31-39.

[4]Thomas Craig K J, Morgan L C, Chen C H, et al. Systematic review of context-aware digital behavior change interventions to improve health. Translational behavioral medicine, 2021, 11(5): 1037-1048.

[5]Scheck, K., Lee, D. Y., Kyung, B. P., Ryu, S. H., Lee, D. L., & Lee, W. B. (2015). An Exercise to Explore Avatar Customization and Gender Swapping. 한국게임학회 논문지, 15(2), 63-71.

축사 환경에서 LSTM 오토인코더 기반 노이즈 감지와 오디오 인페인팅

정세연, 김상철

전북대학교 지능형로봇연구소

e-mail : nuguriie@gmail.com, sckim7777@gmail.com

Noise Detection in Livestock Environments Using LSTM Autoencoders and Audio Inpainting

SeYeon Chung, SangCheol Kim

Core Research Institute of Intelligent Robots, Jeonbuk National University

Keywords: Deep learning, LSTM(Long Short-Term Memory), U-Net, Audio Inpainting

요약

This study investigates the impact of various background noises—such as the operation sound of automatic feeders, animal footsteps, and collision sounds of objects—that occur in the barn environment on the detection and identification of pig vocal signals. These background noises act as major disturbances in the process of collecting and analyzing pig vocal data, and this study proposes a methodology to effectively detect and remove them. In particular, we have precisely detected noise that represents abnormal patterns within the spectrogram images of vocal signals, beyond the anomalies in sensor data occurring in the barn environment, using an autoencoder. Furthermore, we have developed an audio inpainting technique based on the U-Net architecture adopting VGGNet as a feature extractor, and restored the noise-removed vocal signal. Through this, we have successfully derived results that the autoencoder recognizes as normal voices. These research results are expected to contribute to improving the accuracy and reliability of voice-based monitoring systems in barn environments.

1. 서론

이상치 탐지는 정상적인 데이터 패턴에서 벗어난 비정상적인 데이터를 감지하는 알고리즘으로, 시계열 데이터 분석, 패턴 분석 등 다양한 분야에서 활용되고 있다. 본 논문에서는 LSTM(Long Short-Term Memory) 기반 오토인코더를 활용해 축사 환경에서의 이상치 탐지 실험을 제안한다. LSTM은 시간 순서가 중요한 데이터를 처리하며, 장기 의존성 문제를 해결하는 게이트 메커니즘을 가지고 있어 긴 시퀀스 데이터에서도 중요한 정보를 학습할 수 있다.

오토인코더는 입력 데이터를 압축하고 복원하는 과정을 통해 데이터의 주요 특징을 추출하는 비지도 학습 알고리즘이다. 이를 활용하여 축사 환경에서의 센서 데이터를 분석하고, 정상 데이터와 잡음을 구분하는 실험을 수행하였다. 축사 환경의 소리 데이터는 돼지 소리뿐만 아니라 다양한 잡음이 섞여 있어 분석이 큰 어려움이 있다. 기존 연구에서는 잡음을 필터링하는 방식으로 소리를 분류했으나, 본 논문에서는 LSTM 오토인코더를 이용해 소리 데이터를 분석하고 비정상 데이터인 잡음을 감지했다.

이후 잡음이 포함된 오디오 데이터를 무음 처리하고, VGGNet을 피쳐 추출기로 사용하는 U-Net 모델을 통해

오디오를 복원하였다. U-Net는 다양한 해상도의 피쳐 맵을 결합하여 복잡한 패턴과 세부 정보를 잘 포착할 수 있는 구조로, 본 연구에서는 오디오 데이터 복원에 사용되었다.

복원된 오디오 데이터를 다시 LSTM 오토인코더에 입력해 복원의 정확도를 확인하였으며, 이를 통해 본 연구에서 제안한 방법이 축사 환경에서의 오디오 데이터 처리에 효과적임을 검증하였다.

2. 본론

2.1 모델 아키텍처

본 논문에서는 축사 환경에서의 오디오 데이터 처리를 위한 효과적인 방법론을 제안한다. 제안된 모델은 두 가지 주요 단계로 구성되어 있다: 첫 번째 단계는 오토인코더를 활용한 데이터 복원, 두 번째 단계는 오디오 인페인팅이다.

첫 번째 단계에서는 LSTM 기반의 오토인코더를 사용하여 데이터 복원을 수행한다. 오토인코더는 비지도 학습 알고리즘의 한 종류로, 입력 데이터를 압축하고 복원하는 과정을 통해 데이터의 중요한 특징을 학습한다.

두 번째 단계에서는 손상된 오디오 신호를 복원하기 위해 VGGNet을 특징 추출기로 사용하는 오디오 인페인팅

기법을 적용한다. VGGNet은 깊은 층의 구조와 우수한 성능으로 이미지 분류 작업에 널리 사용되는 딥러닝 모델로, 강력한 피쳐 추출 능력을 가지고 있어 중요한 특징을 추출할 수 있다. 이러한 VGGNet의 강력한 피쳐 추출을 이용한 U-Net을 제안한다.

본 논문에서는 LSTM 기반의 오토인코더를 이용하여 음성 의 이상치인 잡음을 감지하고, 딥러닝 기반의 U-Net를 이용하여 개선된 오디오 복원을 수행하였다.

3. 실험 및 성능평가

3.1. 데이터 세트

본 연구에서 사용된 데이터는 국립순천대학교 돼지 농장에서 2022년 9월 7일, 1시간 간격으로 수집되었다. 이를 2초 단위로 세분화하여 총 460개의 오디오 클립으로 구성된 데이터셋을 만들었고, 이를 다시 0.2초 단위로 재분할하여 총 6,100개의 샘플로 구성된 감지 데이터셋을 구축하였다. 데이터셋은 학습, 검증, 테스트 용도로 6:2:2 비율로

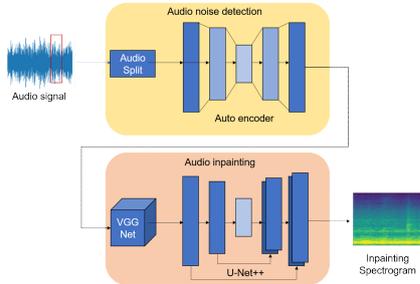


Fig. 1. The overall architecture configuration of the proposed model.

분할되었으며, 멜스펙트로그램으로 변환 후 Min-Max Normalization을 통해 정규화하여 모델 학습에 사용되었다.

3.2. 실험 결과

모델 학습에 사용된 매개 변수는 Epoch 100, Batch size 32, Learning rate는 0.0001, Loss function은 Mean Squared Error(MSE)를 사용하였으며 Optimizer는 ‘Adam’을 사용되었다.

$$Threshold_mean_std = mean_error + 2 \cdot std_error \quad (1)$$

수식 1과 같이 평균 오류와 표준편차를 사용하여 임계값을 설정한다. 이 임계값은 이상치 탐지에 사용된다. 일반적으로 평균값에 표준편차의 몇 배를 더한 값을 임계값으로 설정한다. 본 논문에서는 2배를 사용한다.

이러한 수식으로 임계값 0.2557을 도출하였다. 이와 같은 임계값을 적용함으로써, 다양한 환경적 요인들로 구성된 복잡한 데이터 속에서 노이즈를 효과적으로 감지할 수 있었다.

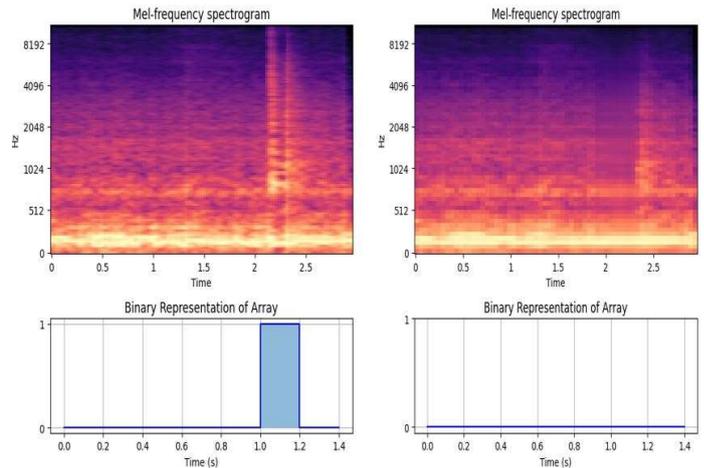


Fig. 2. Noise detection results of the reconstructed audio data Table 1. Performance evaluation results of the LSTM autoencoder model

MSE↓	MAE↓	RMSE↓	R-squared↑
0.0244	0.1249	0.0366	0.6440

Table 2.. Performance evaluation results of the UNet++ model

MSE↓	MAE↓	RMSE↓	R-squared↑
0.0102	0.0221	0.0280	0.6236

3. 결론

본 연구는 국립 순천대학교 돼지사육장에서 수집된 환경 오디오 데이터를 활용해 이상치 감지 및 오디오 인페인팅 모델을 개발하였다. LSTM과 오토인코더를 결합하여 정상 및 비정상 데이터를 효과적으로 분류했으며, VGGNet 기반의 특징 추출기를 사용한 U-Net 오디오 인페인팅 모델을 구현하였다. 실험 결과, 두 모델 모두 낮은 오차 값을 기록하며 복원 능력을 입증하였고, 특히 U-Net 모델의 적용이 성공적으로 오디오 노이즈 식별 성능을 향상시켰음을 확인할 수 있었다.

참고 문헌

[1] Y.-J. Choi and J.-Y. Ryu, "Design of low-power variable gain amplifier," Journal of Institute of Control, Robotics and Systems (in Korean), vol. 28, no. 1, pp. 1-5, Jan. 2022.
doi: <https://doi.org/10.5302/J.ICROS.2022.21.0138>

[2] S. Zhao, J. Wang, H. Xu, and H. Wang, "Finite horizon robust optimal tracking control based on approximate dynamic programming for switched systems with uncertainties," International Journal of Control, Automation and Systems, vol. 20, no. 4, pp.1051-1062, 2022.
doi: <http://dx.doi.org/10.1007/s12555-020-0982-8>

돼지 얼굴 인식을 위한 딥러닝: 종합적인 검출 및 분류 접근법

마리한, 김상철
전북대학교 전자공학부, 전북대학교 지능형 로봇 연구소
e-mail :maruihan1995@gmail.com, sckim7777@gmail.com

Deep Learning for Pig Face Recognition: A Comprehensive Detection and Classification Approach

Ruihan Ma, Sang-Cheol Kim
Jeonbuk National University Department of Electronic and Information
Engineering Intelligenet Robots Research Center, Jeonbuk National University

Abstract

Identifying individual pigs is essential for breeding, health management, and disease control in modern farming. Traditional methods are labor-intensive and inaccurate, while CNN-based pig face recognition models face limitations in computational demands and reliance on extensive labeled data. This paper proposes a novel decoupled approach to pig face recognition, separating detection from identification. A detection model is used as a pre-processing step, reducing the need for extensive re-annotation in new datasets. A small, high-quality dataset of 1500 annotated images from 20 pigs is constructed for detection. The detection model autonomously generates a larger pig face recognition dataset with 56 pig classes. During recognition, ViT model to recognize the pig face. Experimental results demonstrate high detection accuracy on the small dataset and successful generation of a large-scale recognition dataset.

Key Word: Deep Learning; Pig Face Recognition; Face Detection; ViT; Yolo-V8

1. Introduction

With the escalating demands for livestock production and quality, agriculture has rapidly advanced into intensive and smart farming practices aimed at enhancing livestock productivity. In today's world, the advent of new diseases that pose threats to pig health necessitates increasingly vigilant monitoring, and AI technology has been adopted for automated diagnostic tests to facilitate early disease detection in livestock. [1]

Traditional methods for pig identification, such as paint marking and manual observation, have long been used on farms. However, these methods are time-consuming and error-prone, especially with large herds. RFID ear-tagging emerged as an alternative, offering a more streamlined process. RFID systems, though accurate, are costly and raise concerns regarding animal welfare due to the physical implantation required.[2]

To address these limitations, we propose a novel two-stage pig face recognition model that decouples the detection from recognition. This approach reduces the need for repetitive annotation tasks. Once trained, the detection model acts as a pre-processing step, autonomously generating pig face data for recognition without additional human intervention.

In our study, we constructed a small-scale pig face

detection dataset comprising annotated images from 20 pigs. We employed the YOLO-v8 detection model [3] for its efficiency in detecting pig faces. After detection, we applied a Multi-Scale Structural Similarity (MS-SSIM) [4] filtering module to ensure dataset diversity by removing redundant images. Next, the ViT classification model [5] is utilized to identify each pig face.

2. Method

2.1 Dataset Collection, Preparation, and Annotation

Our approach to creating the pig face detection dataset focused on selecting high-quality images from farm-captured footage of 56 pigs. The goal was to optimize the dataset for training an effective detection model while minimizing the annotation workload. To achieve this, we prioritized images with clear visibility of facial features, adequate lighting, and minimal background noise. From the extensive footage, we selected 20 pigs and manually annotated approximately 1,500 high-quality images. The annotated dataset was then split into training, validation, and testing subsets using a 6:2:2 ratio, ensuring a balanced approach for model training and validation. This strategy streamlined data preparation, enhancing the detection model's ability to process and extract relevant information from larger datasets. Such a method proves highly beneficial in smart farming scenarios, where efficient and

accurate data processing is essential.

2.2 Automatic Pig Face Detection

In developing our two-stage pig face recognition model, we implemented a decoupled approach that separates detection and recognition, reducing the need for extensive data annotation. For detection, we utilized YOLO-v8 due to its speed and accuracy, making it ideal for real-time applications. YOLO-v8 uses the CSPDarknet backbone, which improves feature extraction efficiency by reducing parameters and computational load through its Cross-Stage Partial (CSP) structure. The detection head, YOLO-v8-Head, processes and compresses feature maps with convolutional and pooling layers, transforming them into detection results. Detected pig faces are then cropped and passed to the recognition model for identification.

2.3 Pig Face Recognition

In the proposed method, we chose the Vision Transformer (ViT) as our core teacher model for pig face recognition due to its high precision and strong performance. ViT utilizes the self-attention mechanism to process image patches, allowing it to capture global dependencies across the entire image.

3. Experiment

3.1 Implementation Details

To develop an effective pig face detection model, we created a dataset of 1,586 high-quality images from 20 pigs, divided into training, validation, and testing sets with a 6:2:2 ratio. Each image was manually annotated with bounding boxes using the LabelMe tool. This precise annotation enabled robust training on a smaller dataset, which could then be applied to a larger unannotated set for automatic pig face detection.

We fine-tuned four pre-trained detection models from the MS-COCO dataset, using the SGD optimizer with a 0.01 learning rate and advanced data augmentation techniques like Mosaic and RandomFlip. The models were trained for 300 epochs on resized images (224×224 pixels) with a batch size of 64.

3.2 Implementation Result

For training and prediction on our annotated pig face detection dataset, we utilized YOLO-v8 as the sole detection model due to its advanced architecture and superior performance in object detection tasks. YOLO-v8's efficient design makes it well-suited for detecting pig faces with high accuracy, even under challenging conditions. As shown in Table 1, YOLO-v8 achieved outstanding results across key metrics. The model recorded the highest average precision (AP) at IoU thresholds of 0.50 (AP50), 0.75 (AP75), and 0.95 (AP95), with values of 0.990, 0.973, and 0.869, respectively. Additionally, it demonstrated a high recall of 0.895. These results highlight YOLO-v8's excellent capability for accurate pig face detection in diverse scenarios.

Table 1. The result of pig face detection using Yolo-v8.

Model	AP50	AP75	AP50-AP95	Recall
Yolo-v8	0.990	0.973	0.869	0.895



Figure 1. Visualization of detection model prediction using Yolo-v8 on 20 pig face detection dataset

The ViT model demonstrated exceptional performance across all key metrics, achieving an accuracy of 0.972, precision of 0.978, recall of 0.970, and an F1 score of 0.974.

Table2. The result of pig face recognition using ViT model.

Model	Accuracy	Precision	Recall	F1 Score	FLOPs
ViT	0.972	0.978	0.970	0.974	55.5

Acknowledgement

본 논문은 한국연구재단 대학중점연구소지원사업 (NRF-2019R1A6A1A09031717)과 2023년도 정부(농림축산식품부, 과학기술정보통신부, 농촌진흥청 공동)의 재원으로 스마트팜연구개발사업단의 지원을 받아 수행된 연구임(No. 421023-04)

References

- [1] Sihalath, T.; Basak, J.K.; Bhujel, A.; Arulmozhi, E.; Moon, B.E.; Kim, H.T. Pig identification using deep convolutional neural network based on different age range. *J. Biosyst. Eng.* 2021, 46, 182 - 195.
- [2] Rajaraman, V. Radio frequency identification. *Resonance* 2017, 22, 3 - 13.
- [3] Jocher, G.; Chaurasia, A.; Qiu, J. YOLO by Ultralytics. 2023. Available online: <https://github.com/ultralytics/ultralytics>.
- [4] Khalaf, H.A.; Tolba, A.S.; Rashid, M.Z. Event triggered intelligent video recording system using MS-SSIM for smart home security. *AIN Shams Eng. J.* 2020, 9, 1527 - 1533. (accessed on 20 April 2023)
- [5] Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* 2020, arXiv:2010.11929.

생성형 인공지능 기반 기술발전단계 평가 연구

이하윤
한국과학기술정보연구원
미래연구팀
202402007@kisti.re.kr

이주현
한국과학기술정보연구원
미래연구팀
leeju@kisti.re.kr

양혜영
한국과학기술정보연구원
미래연구팀
hyyang@kisti.re.kr

Assessment of Technology Development Stages Using Generative AI

Hayoon Lee
Korea Institute of Science
and Technology
Information

Juhyun Lee
Korea Institute of
Science and Technology
Information

Heyoung Yang
Korea Institute of Science
and Technology
Information

요 약

기술평가는 기술전략과 기술정책 수립 시에 필수적인 근거를 제공하는 연구활동으로, 일반적으로 과학기술분야 빅데이터와 전문가 집단의 지식에 기반하여 연구된다. 최근 생성형 인공지능 GPT가 전문가 임무를 수행하는 영역이 확대되고 있어 기술평가에 대한 활용 가능성을 모색해볼 필요가 있다. 본 연구에서는 GPT-4o에게 바이오분야 기술평가를 위한 전문영역 요청, 해당 전문영역에 대한 전문가 에이전트 생성, 생성된 전문가 에이전트를 활용하여 바이오분야 기술들에 대한 기술발전 단계를 평가하는 실험을 수행하였다. 미래기술 예측과 기술평가에서 많이 활용되는 델파이기법을 응용하여 프롬프트를 구축하였고 총 3라운드에 걸쳐 전문가 에이전트가 기술평가의 수행과 결과 업데이트를 반복하도록 요청하는 방식으로 실험을 구성하였다. 또한 에이전트에게 단순히 질의응답하는 방식의 실험과 “위키피디아에 따르면(according to Wikipedia)”을 추가한 방식의 실험 등 두 가지 실험을 진행하였다. 평가대상의 기술목록은 2012년과 2021년에 발표된 과학기술예측조사의 미래기술 중 바이오분야 기술을 선별하여 활용하였다. 실험결과, 2012년과 2021년 미래기술에 대한 기술발전단계에 대한 평가결과에서 다소 차이가 나타났다.

1. 서 론

미래기술 예측과 기술평가는 과학기술 분야에서 기술전략과 정책 수립의 근간이 되는 중요한 연구 활동이다. 전통적으로 이는 빅데이터와 전문가 지식을 기반으로 수행되어 왔다. 최근 GPT와 같은 생성형 AI의 발전으로 전문가 영역에서의 AI 활용 가능성이 확대되고 있어, 미래기술 예측과 기술평가 분야에서의 GPT 활용 가능성을 탐구할 필요성이 대두되고 있다.

본 연구에서는 GPT-4o를 활용하여 바이오 분야의 기술평가를 수행하는 실험을 진행하였다. 구체적으로, GPT-4o에 바이오 분야 전문가 역할을 부여하고, 이를 통해 생성된 전문가 에이전트를 활용하여 기술발전단계 평가를 수행하였다. 델파이 기법을 응용한 프롬프트를 구성하여 3라운드에 걸친 평가를 진행하였으며, 단순 질의응답 방식의 실험과 "위키피디아에 따르면"이라는 문구를 추가해 위키피디아를 참고하도록 설정한 실험을 수행하였다.

평가 대상은 2012년과 2021년 과학기술예측조사의 바이오분야 미래기술로 선정하였다. 실험 결과, 두 시기의 기술에 대한 평가에서 다소 차이가 나타났으며, 평가에 대한 근거 문장도 최신 기술동향에 기반하여 생성되는 것으로

확인되었다. 본 연구는 GPT-4o를 기술평가에 적용하는 새로운 시도로서, 전통적인 방법론을 보완하고 더 효율적이고 다각적인 기술인텔리전스 체계 구축에 기여할 수 있을 것으로 기대된다. 또한, 인공지능의 활용 범위를 확장하고 과학기술 정책 수립에 새로운 관점을 제공할 수 있을 것이다.

2. 관련 연구: 생성형 AI의 기술평가 및 예측

최근 GPT와 같은 대규모 언어 모델을 활용한 기술 예측 연구가 시작되고 있다. Sidi 등(2023)은 GPT-3를 활용하여 과학기술 논문의 초록을 생성하고, 이를 통해 최신 기술트렌드를 작성하고 미래 연구 동향을 예측하는 방법을 제안하였다. Wang과 Li(2023)는 ChatGPT를 활용하여 특정 기술 분야의 전문가 의견을 시뮬레이션하고, 이를 기반으로 기술 발전 경로를 예측하는 연구를 수행하였다. Zhang 등(2023)은 GPT 기반 모델을 사용하여 특허 문서를 분석하고, 이를 통해 신기술 출현을 예측하는 방법을 제안하였다.

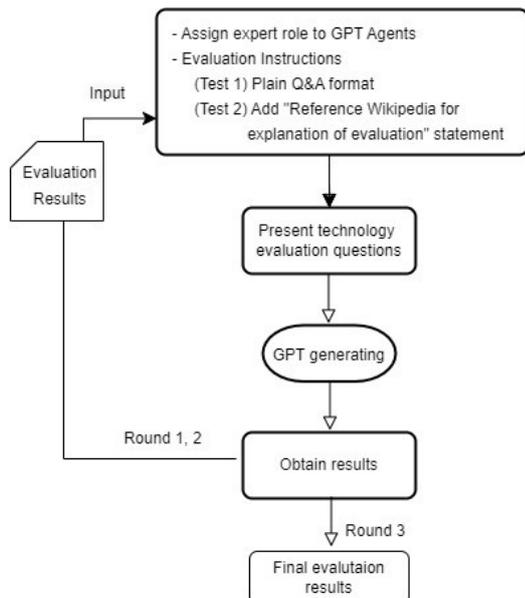
본 연구는 이러한 선행 연구들을 기반으로, GPT-4o를 활용한 기술평가의 새로운 방법론을 제시하고자 한다. 특히,

텔파이 기법의 반복적 합의 도출 과정을 GPT의 에이전트 시스템에 적용함으로써, AI 기반 기술평가방법의 활용 가능성을 모색해보고자 한다.

3. 실험설계

본 연구는 실험1과 실험2로 구성되어있으며, 각 실험은 그림 1과 같은 프롬프트로 구현되었다. 평가에 사용된 바이오분야 기술목록은 2012년 과학기술예측조사 결과 중 기술적, 사회적 실현시기가 빠른 11개 기술, 2021년 과학기술예측조사 결과 중 기술적, 사회적 실현시기가 먼 10개 기술을 활용하였다. 실험1과 실험2는 텔파이기법이 공통적으로 적용되었고, 실험1은 GPT-4o가 생성한 전문가 에이전트들에게 단순 질의응답을 통해 결과를 도출하는 방식으로, 실험2는 “위키피디아에 따르면(Reference Wikipedia for explanation of evaluation)” 방식의 위키피디아 적용을 통해 결과를 도출하는 방식으로 구성되었다. 기술발전단계는 총 7단계로 구분하여 답변하도록 하였다. 기술발전단계의 구분은 표 1과 같다.

(그림 1) 실험1과 실험2의 프롬프트 구성



(표 1) 기술발전단계의 7단계

1	Long time is needed until technical realization
2	Technical realization is not far off
3	Technically realized
4	Early stage of social commercialization
5	Stage of expanding commercialization
6	Stage where the technology market is saturated
7	Stage where the technology market is shrinking

4. 실험결과

실험1, 2의 결과는 표 2와 같다. 실험1과 실험2에서 전체적으로 2012년 미래기술이 2021년 미래기술보다 기술발전단계가 다소 높게 나타났다. 그러나 그 차이가 매우 작아 GPT 에이전트들이 기술트렌드 정보를 평가 시 활용하기

는 하나 충분한 전문적 지식에 기반한다고 보기는 어렵다. 또한 실험1보다는 실험2의 결과가 일관성을 보여주고 있어, “위키피디아에 따르면” 프롬프트가 GPT 에이전트들의 전문성을 확보하는데 도움을 주는 것으로 판단된다.

(표 2) 실험1, 2의 기술발전단계 평가결과

실험1		기술별 전문가 평가 평균값 변화										
2012년 미래기술	1라운드	5.0	4.0	4.0	4.0	3.9	4.1	4.2	3.9	3.6	4.1	3.8
	2라운드	5.0	4.1	4.0	4.1	4.0	4.0	4.3	4.0	4.0	4.1	4.0
	3라운드	5.0	1.0	4.1	4.0	4.0	4.0	2.0	4.0	4.0	4.0	4.0
2021년 미래기술	1라운드	1.2	3.0	2.0	4.0	3.1	3.9	3.4	4.0	2.0	3.9	
	2라운드	1.0	4.0	2.0	4.0	4.0	4.0	4.0	4.0	2.5	4.0	
	3라운드	4.6	4.0	4.0	4.0	4.0	4.0	4.0	4.0	3.0	4.0	4.0
실험2		기술별 전문가 평가 평균값 변화										
2012년 미래기술	1라운드	4.9	4.0	4.0	3.9	3.8	4.2	4.3	4.0	3.7	3.9	3.1
	2라운드	5.0	4.0	4.0	4.0	4.0	4.6	4.3	4.0	4.0	4.1	4.0
	3라운드	5.0	4.0	4.0	4.0	4.0	5.0	4.8	4.0	4.0	4.0	4.0
2021년 미래기술	1라운드	1.2	3.2	2.0	4.0	3.5	3.9	3.4	4.0	2.3	3.9	
	2라운드	1.0	4.0	2.0	4.0	4.0	4.0	4.0	4.0	2.5	4.0	
	3라운드	1.0	4.0	2.0	4.0	4.0	4.0	4.0	4.0	2.8	4.0	

5. 결론

본 연구는 생성형 인공지능 GPT-4o를 활용한 기술발전단계 평가의 가능성을 탐색적으로 살펴보았다. 연구 결과를 바탕으로 다음과 같은 결론을 도출할 수 있다.

첫째, GPT-4o와 텔파이기법을 응용한 기술발전단계 평가는 기존의 전문가 기반 평가 방식을 보완할 수 있는 잠재력을 보여주었다.

둘째, 2012년과 2021년 미래기술에 대한 평가 결과의 차이는 GPT-4o가 시간의 흐름에 따른 기술 발전을 어느 정도 반영할 수 있음을 보여준다.

셋째, "위키피디아에 따르면"이라는 문구를 추가한 실험 결과는 GPT-4o의 답변에 신뢰성과 객관성을 부여하는 데 일정 부분 효과가 있음을 보여주었다.

그러나 본 연구는 다음과 같은 한계점을 가지고 있다. 2012년과 2021년 미래기술에 대한 평가 결과의 시간적 차이가 명확하지 않다. 이는 기술에 대한 상세설명이 프롬프트로 입력되지 않았기 때문으로 해석되어 추후 보완이 필요하다. 또한, GPT-4o의 평가 결과를 실제 전문가 평가와 직접적으로 비교하지 않아, 그 정확성과 신뢰성을 완전히 검증하지 못했다는 점도 한계로 지적될 수 있다.

참고문헌

[1] Sidi, F., Jabar, M. A., Mustapha, A., Sani, N. F. M., & Ishak, I. (2023). Generating scientific paper abstracts using GPT-3 for predicting future research trends.
 [2] Wang, X., & Li, F. (2023). ChatGPT-based expert opinion simulation for technology foresight: A case study in the field of brain-computer interfaces.
 [3] Zhang, Y., Wu, M., Liu, H., & Shen, L. (2023). GPT-based patent analysis for emerging technology identification: A case study in the field of blockchain.

군집화와 객체 인식을 이용한 전방위 융합 객체 인식 시스템

김동훈¹, 조세운¹, 김병준¹, 김서정¹, 조연호¹, 현상미¹, 배성재¹, 최경호², 최나영², *정성환¹
¹한국전자기술연구원, ²에코스ENC

e-mail : clickmiss123@keti.re.kr, swcho@keti.re.kr, jun0420@keti.re.kr, scott3554@keti.re.kr, geijin0821@keti.re.kr, tkdal6365@keti.re.kr, bsj940528@keti.re.kr, donaldkho@naver.com, ani1717@naver.com, *shjeong@keti.re.kr

Omnidirectional Fusion Object Recognition System Using Clustering and Object Detection

¹Donghoon Kim, ¹SeWoon Cho, ¹Byoungjun Kim, ¹Seojeong Kim, ¹Yeonho Jo, ¹Sangmi Hyeon, ¹Seongjae Bae, ²Kyeongho Choi, ²Nayoung Choi, ¹Shunghwan Jeong*
¹Korea Electronics Technology Institute, ²EKOSENC

요약

본 논문에서는 카메라와 라이다 센서를 이용하여 주변 전방위에서 나타나는 객체 정보를 융합하여 모니터링을 수행하는 시스템을 설계하였다. 시스템은 센서의 배치와 센서 간 캘리브레이션 파라미터의 생성, 알고리즘을 통한 센싱 정보의 융합으로 구성된다. 실험에서 전방위 센싱을 위해 고정기구를 이용하여 다수 개 카메라와 하나의 라이다 센서를 배치하고 각각에 대한 캘리브레이션을 수행하였다. 또한 적응적 클러스터링을 이용한 포인트 클라우드 군집 검출, YOLOv5 모델을 이용한 영상 객체 검출을 수행함으로써 실시간으로 인식 정보를 생성할 수 있도록 구성하였다. 전체 시스템은 센서별 생성된 군집 및 객체 인식 정보와 센서 간 캘리브레이션 정보를 기반으로 하여 클러스터링 결과에 존재하지 않는 클래스 정보, 객체 인식 결과에 존재하지 않는 거리 정보를 상호 교환하는 형태로 인식 정보의 융합을 수행한다.

주요어 : 전방위 감지, 클러스터링, 객체 인식, 캘리브레이션, 인식 결과 융합

1. 서론

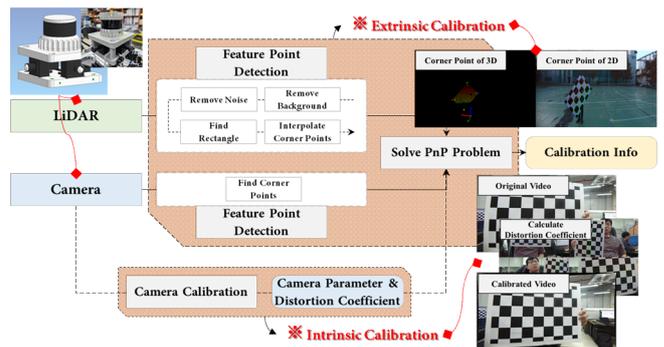
차량과 같은 구동체의 이동에 대한 안전사고 방지는 항상 중요한 문제로 다루어져 왔다. 이러한 안전사고 방지는 센서와 데이터 처리 기술의 발전에 힘입어 빠른 성장 중인 자율 주행 기술로 인해, 보다 중요 해지고 있다. 이러한 안전의 확보를 위해 센서를 통한 주변 환경의 스캔과 이에 대한 처리의 수행이 필요하게 된다. 센서는 타입에 따라 장단점이 존재하는 다양한 데이터가 발생하며 이러한 데이터 특성은 처리 결과에 영향을 미치게 된다. 본 논문에서는 이러한 센서 타입별 특성을 센서 인식 결과의 융합을 통해 결합하기 위한 연구를 진행하였다.

카메라와 라이다 센서 사이의 관계는 캘리브레이션을 통해 얻을 수 있는 외부 파라미터로 수치적 표현이 가능하며 이를 활용하면 센서 데이터 상호 간의 변환이 가능하다. 본 논문에서는 데이터 간 상호 변환 가능한 특성으로부터 상호 참조로 센서의 단점을 줄일 수 있는 시스템을 설계하였다. 이를 위해 카메라와 라이다 사이의 캘리브레이션을 수행하여 외부 파라미터를 계산하였고, 영상에 대한 객체 인식과 포인트 클라우드에 대한 클러스터링을 각각 수행하였다. 마지막으로 클러스터링 및 객체 인식 결과에 대해 외부 파라미터를 적용함으로써 객체-군집 사이의 관계를 확인하여 각 데이터에 존재하지 않는 거리와

분류 정보를 참조할 수 있도록 시스템을 구성하였다.

2. 본론

2.1 카메라-라이다 캘리브레이션

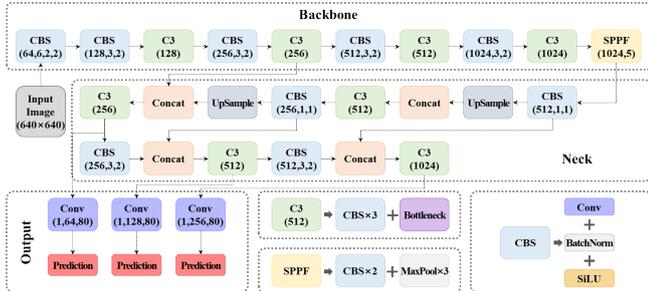


(그림 1) 카메라-라이다 캘리브레이션 흐름도

본 논문에서는 ‘Ouster OS1-64’ 라이다, ‘ELP 4K HDR usb’ 카메라, 2.8GHz CPU, NVIDIA RTX 2060 GPU를 이용하여 실험을 진행하였다. 전방위의 감지를 위한 4개의 카메라를 배치하였으며, 그림 1과 같은 알고리즘으로 각각의 카메라와 라이다 사이의 캘리브레이션을 수행하였다. 라이다의 낮은 수평 해상도로 인해 거리가 멀어질수록 캘리브레이션 보드의 정확한 형태 추정이 어려워져 후보정 과정[1]이 요구되기 때문에 영상-포인트 클라우드 지점 매칭은 1~5m 이내의 데이터로 제한하였다.

2.2 영상 객체 인식과 포인트 클라우드 클러스터링

본 연구에서는 단일 PC에서 융합 전 단계가 실시간으로 수행될 수 있도록 객체 검출과 클러스터링에 빠른 실행이 보장되는 모델을 사용하였다. 이에 따라 영상 데이터에서의 객체 검출에는 MS-COCO 데이터셋으로 학습된 그림 2와 같은 구조의 YOLOv5 모델[2]을 사용하였다.



(그림 2) 2D 객체 검출에 사용한 YOLOv5 모델 구조

실험 환경에서의 실시간 처리를 위해 스레드 내에서 객체 검출 모델의 테스트를 진행하였고, 네트워크 규모에 따라 동시 4개 영상 데이터를 입력으로 하여 표 1과 같은 처리 속도를 확인하였다. 실험으로부터 라이다 센서 초당 프레임 수인 10FPS 이상의 처리 속도를 가지며 가장 검출 성능이 높은 Large 규모의 네트워크를 선택하였다.

(표 1) 실험 환경에서 네트워크 규모에 따른 프레임 처리 속도

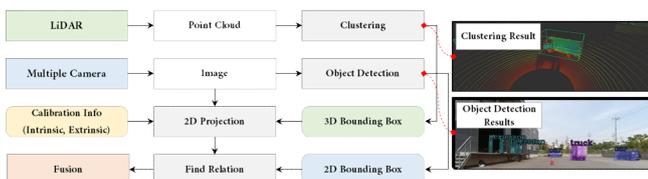
	YOLOv5 네트워크 규모				
	Nano	Small	Medium	Large	XLarge
FPS	44.32	30.05	19.98	10.87	6.98

포인트 클라우드 데이터의 경우 분류되지 않은 객체를 찾아냈음을 가정하여 클러스터링 모델을 이용한 군집을 검출하였다. 라이다의 포인트 클라우드 데이터는 거리가 멀어지는 데 따라 밀집도가 점차 떨어져 일정한 임계값으로 군집화를 수행하기 어려울 수 있다. 따라서 센서로부터 원형으로 나타나는 데이터 분포에 대해 거리에 따라 영역을 원형으로 나누고 클러스터링 임계값 파라미터를 차등 적용하는 적응적 클러스터링[3] 방법을 사용하였다.

2.3 융합 객체 인식

포인트 클라우드 데이터의 클러스터링과 영상 데이터의 객체 인식 결과로 각각 3D 바운딩 박스와 2D 바운딩 박스를 얻을 수 있다. 또한, 앞서 추정된 캘리브레이션 정보인 카메라 초점 거리 f 와 렌즈 중심 c , 카메라 좌표계와 월드 좌표계 사이의 회전 r , 평행이동 t 에 대해 수식 1을 통한 3D 데이터의 2D 평면 투영이 가능하다.

$$S \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (\text{수식 1})$$



(그림 3) 융합 객체인식 시스템 흐름도

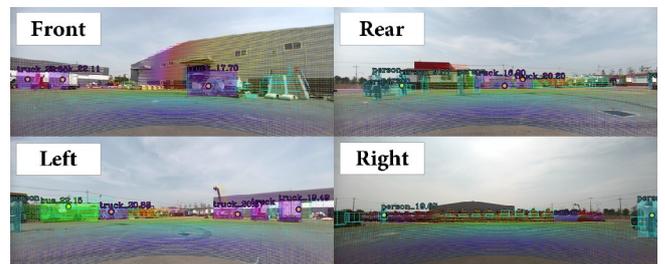
본 논문에서는 3D 바운딩 박스를 2차원 평면에 투영하

고, 투영된 3D 바운딩 박스와 2D 바운딩 박스 사이의 관계를 설정함으로써 그림 3과 같은 구조로 융합 객체 인식 시스템을 구성하였다.

하나의 영상에 대해 융합 인식을 수행한 결과 그림 4와 같이 2D 객체의 거리와 3D 객체의 클래스를 추정할 수 있음을 확인하였으며, 그림 5와 같이 전방위 융합 객체 인식 시스템의 실행을 확인하였다.



(그림 4) 2D 객체인식과 3D 클러스터링의 융합 결과



(그림 5) 전방위에 대한 실시간 융합 객체 인식 결과

3. 결론 및 향후 연구

본 논문에서는 2D 객체인식 결과와 3D 데이터의 군집에 대해 캘리브레이션 정보를 이용한 융합을 수행함으로써 센서 간 정보 교환을 수행하는 시스템을 제안하였다. 시스템으로부터 각 센서에 부족한 정보를 생성할 수 있음을 확인하였다. 본 시스템은 2D 데이터의 3차원 공간상 위치 정보를 확인할 수 없다는 한계가 있으며, 추후 이를 극복하기 위한 연구를 진행할 예정이다.

※Acknowledgement

- 이 연구는 2024년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(20022634, 이동식 유무선 전기차 자동충전시스템 요소부품 및 운용기술 개발)

참고 문헌

- [1] Zhou, Lipu, Zimo Li, and Michael Kaess, "Automatic extrinsic calibration of a camera and a 3d lidar using line and plane correspondences", 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018.
- [2] Kim, Jun-Hwa, et al., "Object detection and classification based on YOLO-V5 with improved maritime dataset", Journal of Marine Science and Engineering. Vo.10 No.3, 2022.
- [3] Yan, Zhi, Tom Duckett, and Nicola Bellotto, "Online learning for 3D LiDAR-based human detection: experimental analysis of point cloud clustering and classification methods", Autonomous Robots. Vol.44 No.2, 2020. pp.147-164.

RTK GNSS와 함께 Dual Extended Kalman 필터를 사용하는 자율 항법 시스템

알리하산[1,2], 김상철[*],2]

[1]전북대학교 전자정보공학부

[2]전북대학교 지능형로봇연구소

e-mail : 202350496@jbnu.ac.kr, *sckim7777@gmail.com

Autonomous Navigation System Using Dual Extended Kalman Filter with RTK GNSS

Hassan Ali[1,2], SangCheol Kim[2]

[1]Department of Electronics and Information Engineering, Jeonbuk National University

[2]Core Research Institute of Intelligent Robots, Jeonbuk National University

Abstract

This paper explores the method of autonomous navigation of a ground vehicle using a two-level Extended Kalman filter application. Vehicle's wheel odometry, magnetic heading, an IMU unit and a high-precision GNSS source with RTK enabled are the sole sensory inputs used. Despite providing highly precise data, the RTK GNSS provides discrete data and is prone to outages which can cause the system to diverge indefinitely. To solve this, the two-level EKF solution is used to interpolate the missing data using vehicle's orientation and velocities as seen by the wheel odometry and ground velocity from a global perspective. The system is implemented on ROS Melodic, achieving a net RMSE of 17.9 cm between ground truth way-points and the resulting trajectory obtained after following the way-points.

1. Introduction

With the rise in Automation industry and Robotics, research in self-driving vehicles has seen many breakthroughs uncovering a number of state-of-the-art methods. Kalman filter is widely used for sensor data fusion as well as state-estimation factoring in probabilistic noise. However, Kalman filter only works well with linear dynamic systems where both the measurement and motion models are linear, whereas in reality a system might be subject to non-linearity which if not dealt with can cause the system's error to grow unbounded. For this purpose, Extended Kalman Filters (EKF) are employed in such complex cases. Instead of using a linear system of equation for state, differentiable functions can be used, following are the characteristic equations of the Extended Kalman Filter:

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}) + w_{k-1} \\ z_k &= h(x_k) + v_k \end{aligned}$$

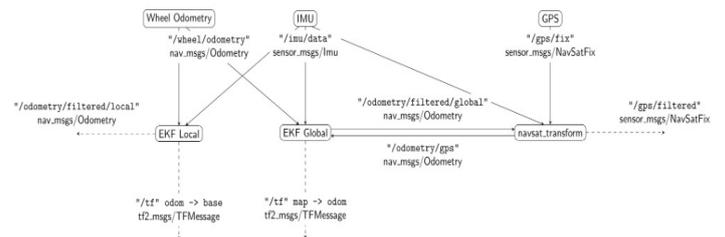
w_k and v_k are process and measurement noises respectively, determined by their corresponding covariance matrices.

2. Methodology

2.1 Working

Two instances of EKF are utilized for our implementation

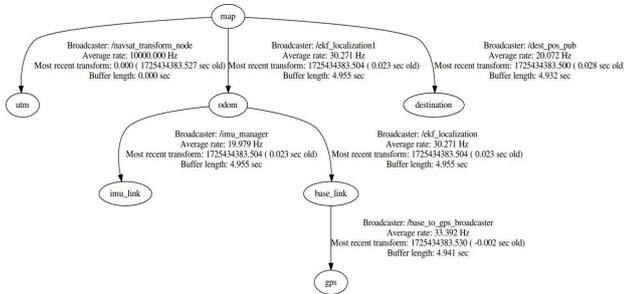
i.e., Local EKF and Global EKF. It is not possible to fuse GPS data directly with the IMU and wheel odometry as all these measurements have a completely different acquisition rate, secondly for the system to perform smoothly, there must be correspondance between the local and global information. Following figure shows the inputs to local and global EKF instances and how they interact.



(Figure 1) Inputs to EKF instances[1]

Both the EKF instances take encoder odometry and IMU as input. The local EKF instance only takes the aforementioned input sources and fuses them, the resulting fused data provides a precise local position and orientation as it predicts the best relative position and orientation of the robot based on its starting point. The next step is to establish a correspondance between the robot's global and relative (local) position, knowing that global EKF fuses the same "local" data along

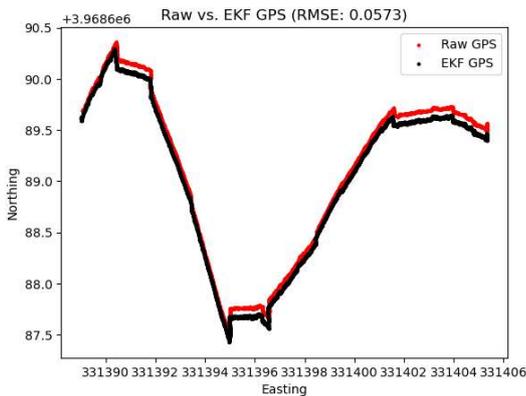
with GPS data, this EKF periodically corrects the local EKF error in order to reduce the long term errors. This suggests the local reference frame used by the first EKF is highly dynamic from the global EKF perspective, because as soon as the global EKF updates the local information is also updated. Figure 2 illustrates the whole process:



(Figure 2) Interaction of different frames and their rate of update. Notice how the global frame “map” transforms to local frame “odom” which transforms to the robot frame “base_link”.

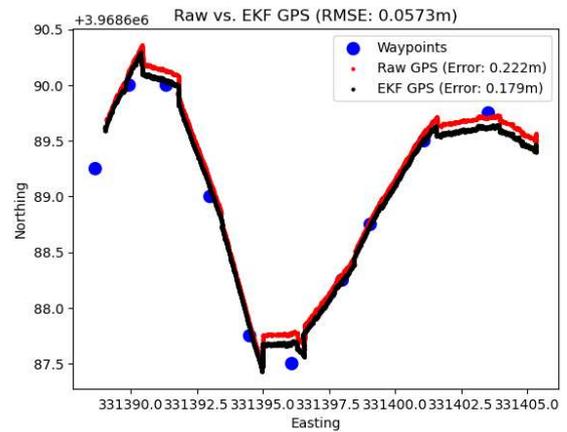
2.2 Results:

The following figure shows the comparison between the raw GPS values obtained directly from the device and the EKF corrected GPS data. Since the device is RTK enabled, the difference between the trajectories is quite small i.e., 5.73 cm.



(Figure 3) Raw vs. EKF GPS

The figure 3 shows the trajectory resulting from a set of predefined way-points. Figure 4. shows the way-points and the mean error between the ground truth trajectory set by the way-points and the resulting trajectories from raw and EKF based GPS data.



(Figure 4) Raw vs. EKF GPS Error with respect to ground-truth trajectory

The EKF corrected GPS trajectory is much closer to the ground-truth as compared to that of raw GPS.

Acknowledgement

This research was supported by the Regional Innovation Strategy (RIS) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2023RIS-008), the Agricultural Policy Support Technology Development Project funded by the Rural Development Administration’s National Institute of Crop Science (RS-2023-00232338), and the University-Focused Research Institute Support Program funded by the National Research Foundation of Korea (NRF-2019R1A6A1A09031717).

References

[1] Moore, Thomas, and Daniel Stouch. "A generalized extended kalman filter implementation for the robot operating system." Intelligent Autonomous Systems 13: Proceedings of the 13th International Conference IAS-13. Springer International Publishing, 2016.

[2] Akhlaghi, Shahrokh, Ning Zhou, and Zhenyu Huang. "Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation." 2017 IEEE power & energy society general meeting. IEEE, 2017.

이차전지 산업의 최상위 수출입 네트워크 분석

박 완, 전홍우, 권이남
 한국과학기술정보연구원 미래기술분석센터 GVC 연구팀
 e-mail : {spring19, hw.chun, ynkwon}@kisti.re.kr

An Analysis of the Top Export-Import Trade Networks in the Secondary Battery Industry

Wan Park, Hong-Woo Chun, Lee-Nam Kwon
 GVC Research Team, Future Technology Analysis Center, Korea Institute of Science and Technology Information

요 약

본 연구는 2013년부터 2023년까지 최근 10년간 수출입 무역데이터를 바탕으로 이차전지 산업에서 주요 6개국(한국, 중국, 일본, 미국, 러시아, EU)의 수출입 네트워크를 분석하였다. 연구 결과, 한국은 2019~2020년에 수출이 급증한 반면, 미국과 EU는 2021년 이후 감소세를 보였다. 수출 네트워크에서는 미국, 중국, 한국 간의 연결성이 강화되었고, 수입 네트워크에서는 인도네시아가 주요 허브로 부상하였다. 특히 인도네시아와 말레이시아는 이차전지 원료를 보유하여 글로벌 공급망에서 중요한 역할을 담당하고 있음을 확인하였다.

1. 서 론

국제무역은 국가 간 상호작용의 주요 통로로, 한 국가에서 발생한 경제적 변화가 네트워크를 통해 전 세계로 확산될 수 있다[1]. 2019년부터 심화된 미·중 무역 갈등으로 공급망 위기는 우리나라의 핵심 수출산업 중 하나인 이차전지 산업에도 영향을 미쳤다. 특히 우리나라는 무역 의존도가 높은 경제 구조를 가지고 있기 때문에, 이차전지 산업의 네트워크 특성을 파악하고 그 특성들을 분석하는 것은 향후 공급망 위기에 대응하는 데 중요한 자료가 될 것이다.

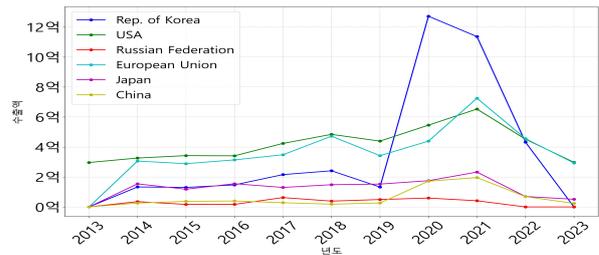
2. 연구 범위 및 방법

본 연구는 이차전지 산업과 관련된 총 39개의 HS 코드를 대상으로, 2013년부터 2023년까지 주요 6개국(한국, 중국, 일본, 미국, 러시아, EU)의 수출입 데이터를 각각 추출하여 분석하였다. 선행 연구에 따르면, 최상위 관계에 기반한 무역 네트워크는 국가가 무역 상대국을 선택할 때의 선호도를 반영할 수 있다[2]. 따라서 본 연구에서도 수출입 상위 1위 및 2위 국가를 대상으로 네트워크 분석을 진행하였다.

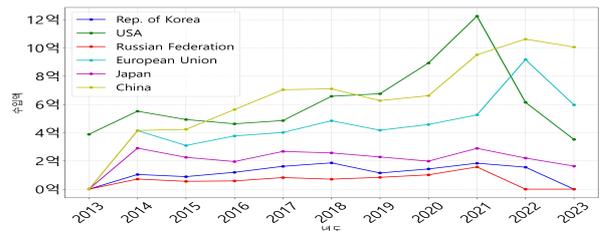
3. 연구 결과

3.1 이차전지 산업의 년도별 수출입 현황

그림 1과 그림 2는 연도별 수출입 현황을 보여준다. 그림 1에서, 한국은 2019년과 2020년에 수출이 두드러지게 상승했으며, 미국과 EU는 일정한 상승세를 보이다가 2021년을 기점으로 감소세로 돌아섰다. 러시아, 일본, 중국은 큰 변동이 없었다. 그림 2의 연도별 수입 현황을 살펴보면, 미국과 중국의 수입이 2019년 이후 급격히 증가한 점이 주목할 만하다.



(그림 1) 이차전지 산업 품목의 국가별 연도별 수출 현황

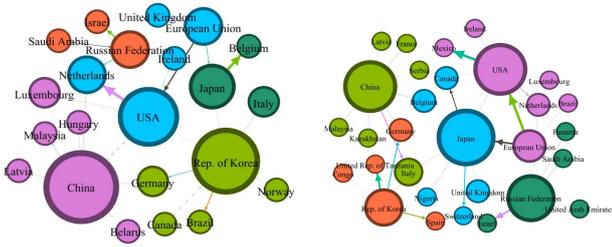


(그림 2) 이차전지 산업 품목의 국가별 연도별 수입 현황

3.2 이차전지 산업의 최상위 수출입 무역 네트워크

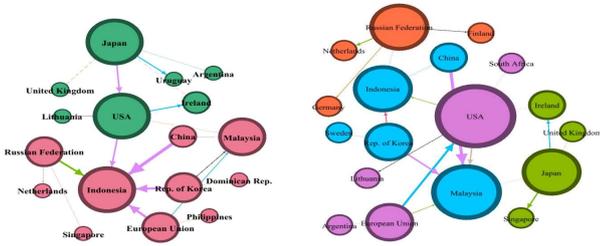
그림 3과 그림 4는 주요 6개국의 수출 상위 1위 및 2위 국가들의 네트워크를 보여준다. 노드(원)의 크기와 선(화살표)의 굵기는 연결 정도(degree)를 나타내며, 노드의 색상은 노드의 연결 수에 따라 형성된 집단(modularity)을 의미한다. 선(화살표)의 색상은 2013년부터 2023년까지의 연도별 변화를 나타낸다. 그림 3을 보면, 수출 1위 국가 네트워크에서는 중국, 한국, 미국, 일본이 비교적 큰 노드를 가지고 있으며, 미국과 모두 밀접하게 연결되어 있다. 러시아와 EU는 미국보다는 다른 국가들과의 연결성이 크다. 최근 몇 년 동안 미국, 중국, 한국 사이의 연결성이 강해진 것을 알 수 있다. 수출 2위 국가 네트워크에서는 중국을 중심으로 구성되어 있으며, 미국, 일본, 한국이 중국

과 밀접하게 연결되어 있다. 또한, 2020년대 초반부터 중국과 미국, EU 사이의 수출 네트워크가 강화된 것을 엿지 색상으로 확인할 수 있다. (핑크색: 2022년)



(그림 3) 수출 상위 1위(좌) 및 2위(우) 국가 네트워크

그림 4에서 주요 6개국의 수입 상위 1위 국가 네트워크를 살펴보면, 인도네시아가 네트워크 중심에 위치하며 가장 큰 노드로 여러 국가와 강하게 연결되어 있다. 이는 인도네시아가 주요 수입 허브 역할을 하고 있음을 나타낸다. 특히 인도네시아, 미국, 중국 간의 연결이 최근 2~3년 동안 강화된 것을 알 수 있다. 수입 상위 2위 국가 네트워크에서는 미국이 가장 큰 노드를 차지하며, 이는 전 세계적으로 가장 많은 국가와 무역 관계를 맺고 있다는 것을 보여준다. 인도네시아, 말레이시아, 중국이 미국과 강하게 연결되어 있으며, EU와 일본도 주요 파트너로 나타난다. 1위 국가 네트워크와 마찬가지로, 미국과 인도네시아, 말레이시아, 중국 간의 연결이 최근 몇 년 동안 강화된 것으로 보인다.



(그림 4) 수입 상위 1위(좌) 및 2위(우) 국가 네트워크

3.3 이차전지 산업의 년도별 최상위 수출입 국가

표 1과 표 2는 주요 6개국이 이차전지 산업의 수출액과 수입액에서 어떤 국가와 주로 교류하고 있으며, 시간이 지남에 따라 교역 상대국이 어떻게 변해왔는지를 보여준다. 한국은 네덜란드, 이스라엘, 사우디아라비아가 수출 상위 1위 국가로 변해왔고, 일본은 2015년부터 2021년까지 벨기에가 수출 1위 국가였으나, 2022년에는 한국으로 바뀌었다. 미국은 네덜란드가 꾸준히 수출 1위를 차지하며 안정적인 파트너십을 유지하고 있다. 수입 상위 1위 국가의 변화를 보면, 주요 6개국 대부분이 인도네시아와 말레이시아를 주요 수입국으로 두고 있는데, 이는 이들 국가가 이차전지의 핵심 광물인 니켈과 리튬을 다수 보유하고 있기 때문으로, 이들 국가가 이차전지 산업과 관련된 글로벌 공급망에서 중요한 역할을 하고 있음을 알 수 있다.

(표 1) 년도별 주요 6개국의 수출 상위 1위 국가 변화

년도	한국	일본	중국	미국	EU	러시아
2013				네덜란드		
2014	네덜란드	이탈리아	말레이시아	네덜란드	일본	네덜란드
2015	이스라엘	벨기에	독일	네덜란드	미국	이스라엘
2016	이스라엘	벨기에	말레이시아	네덜란드	미국	이스라엘
2017	사우디	벨기에	미국	네덜란드	일본	사우디
2018	이스라엘	벨기에	벨라루스	룩셈부르크	미국	이스라엘
2019	사우디	벨기에	네덜란드	네덜란드	미국	사우디
2020	네덜란드	벨기에	헝가리	아일랜드	영국	네덜란드
2021	이스라엘	벨기에	라트비아	네덜란드	영국	이스라엘
2022	노르웨이	한국	룩셈부르크	네덜란드	영국	
2023		러시아	러시아	네덜란드	영국	

(표 2) 년도별 주요 6개국의 수입 상위 1위 국가 변화

년도	한국	일본	중국	미국	EU	러시아
2013		이르랜드		아일랜드		
2014	말레이시아	우루과이	인도네시아	말레이시아	필리핀	인도네시아
2015	말레이시아	우루과이	말레이시아	아일랜드	말레이시아	인도네시아
2016	인도네시아	우루과이	인도네시아	아일랜드	말레이시아	인도네시아
2017	인도네시아	아일랜드	인도네시아	인도네시아	인도네시아	인도네시아
2018	도미니카	영국	인도네시아	인도네시아	인도네시아	인도네시아
2019	인도네시아	미국	인도네시아	인도네시아	인도네시아	네덜란드
2020	인도네시아	미국	인도네시아	리투아니아	인도네시아	네덜란드
2021	인도네시아	미국	인도네시아	리투아니아	인도네시아	싱가포르
2022	인도네시아	미국	인도네시아	인도네시아	인도네시아	
2023		미국	인도네시아	인도네시아	인도네시아	

4. 결 론

본 연구는 최근 10년간 이차전지 산업의 수출입 네트워크를 분석하였다. 분석 결과 한국, 미국, 중국이 이차전지 산업 무역에서 주요 역할을 하고 있으며, 인도네시아와 말레이시아가 핵심 자원 공급국으로서 중요한 위치를 차지하고 있음을 확인했다. 2019-2020년에 한국의 수출이 급증한 반면, 미국과 EU는 2021년 이후 감소했다. 본 연구는 이차전지 글로벌 공급망에 대한 이해를 바탕으로 공급망 위기에 대응하는 자료를 제공하며, 무역 네트워크 변화 분석을 통해 향후 정책 및 전략 수립에 기여할 것이다. 자원 부국과의 무역 강화는 원자재 확보 전략에 도움을 줄 수 있다. 향후 연구에서는 더 많은 국가를 포함하여 이차전지 산업의 글로벌 무역 네트워크를 보다 포괄적으로 분석할 필요가 있다. 또한, 이차전지 산업의 장기적인 무역 네트워크 변화 추세를 예측하는 연구도 필요하다.

※이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업 및 자체사업으로 수행된 연구입니다. (기본사업 과제번호: (KIST)DK24L3MIC2, 자체사업 과제번호: (KIST)J24JR044-24)

참 고 문 헌

- [1] 임병학. “유럽 국가들 간 무역 네트워크 영향력과 핵심주변 분석을 통한 국가 충격의 전염효과에 대한 연구”, 『유라시아연구』 9(1), 2012, pp.93-117.
- [2] Zhou, Min, Wu, Gang and Xu, Helian, “Structure and formation of top networks in international trade, 2001-2010”, Social Networks, 44, 9-21, 2016

웹기반 Google OTP 관리 시스템 구축

김애숙
한남대학교 교양학부
e-mail : 60203@m365.hnu.ac.kr

Building a web-based google OTP management System

Ae-Sook Kim
Talmage Liberal Arts College, Hannam University

요 약

본 연구는 Google OTP(One-Time Password) 기반 2단계 인증 시스템의 효과적인 관리를 위한 웹 기반 프로그램 개발을 목표로 한다. 최근 사이버 보안 위협이 증가함에 따라, 2단계 인증의 중요성은 날로 커지고 있으나, 이를 효율적으로 관리하기 위한 도구는 부족한 실정이다. 본 논문에서는 파이썬 Flask와 nodejs를 이용하여 사용자 친화적인 인터페이스와 서비스에 필수적인 기능을 제공하는 웹 기반 Google OTP 관리 프로그램을 제안한다. 이를 위해 최신 웹 기술과 보안 프로토콜을 사용하여 시스템을 설계하였으며, 주요 기능으로는 OTP 발급, 재발급 및 인증코드 테스트와 사용자 계정 연동 기능을 포함한다. 개발된 프로그램은 서비스 관리자들이 더욱 편리하게 OTP 관리에 기여하며, 향후 연구로는 다양한 인증 시스템과의 통합 및 추가 보안 기능 개발을 제안한다.

1. 서 론

오늘날 사이버 보안 위협이 급증하면서, 개인 정보와 시스템 보안을 보호하기 위한 강력한 인증 메커니즘이 필수적이다. 그중에서도 2단계 인증(2FA)은 사용자 계정을 보호하는 효과적인 방법으로 널리 사용되고 있다. 2단계 인증은 단일 비밀번호 기반의 인증보다 보안성이 높아, 금융, 의료, 공공 서비스 등 다양한 분야에서 빠르게 채택되고 있다. Google Authenticator와 같은 OTP(One-Time Password) 기반의 2단계 인증 시스템은 사용자에게 추가적인 보호 계층을 제공하며, 비밀번호 유출이나 피싱 공격에 대한 강력한 방어 수단이 된다.

그러나, 많은 관리자가 OTP 기반 2단계 인증을 도입하면서 어려움을 겪고 있다. 특히 다수의 계정과 서비스를 관리하는 환경에서는 OTP 관리의 복잡성이 높아지며, 이를 효율적으로 관리할 수 있는 도구의 부재는 시스템 운영과 보안 유지에 큰 도전 과제가 된다. 현재 대부분의 2단계 인증 시스템은 사용자 단위의 개별 설정과 관리가 필요하며, 여러 계정을 관리하는 데 있어 불편함을 초래하고 있다. 이러한 이유로 대규모 서비스에 OTP를 적용하는 경우, 별도의 관리 도구를 제공하는 상용 OTP를 적용하여 2단계 인증을 적용한다.

본 논문은 Google OTP와 같은 오픈 소스 2단계 인증 시스템을 효율적으로 관리하는 방법을 제시함으로써, 사이버 보안의 핵심 요소인 인증 절차의 보안성과 편의성을 동시에 강화하는 데 이바지할 것이다. 이를 통해 조직과 개인이 직면한 보안 위협을 줄이고, 안전한 디지털 환경을 구축하는 데 중요한 역할을 할 것으로 기대한다.

2. 관련 연구

2.1 Google OTP(Google Authenticator)

TOTP(Time-based One-Time Password) 알고리즘을 기반으로 하는 대표적인 2단계 인증 앱으로, RFC 6238 표준에 따라 일정 시간마다 변경되는 일회용 비밀번호를 생성한다. 기존의 HOTP(HMAC-based One-Time Password) 기반 시스템은 카운터를 기준으로 OTP를 생성하지만, TOTP는 시간을 기준으로 하여 더 높은 보안성을 제공한다. 이와 같은 TOTP 기반 시스템은 물리적 토 큰 없이도 안전하고 간편하게 다중 계정 인증을 관리할 수 있어, 개인 사용자뿐만 아니라 기업 환경에서도 널리 채택되고 있다.

2.2 PyOTP

Python 프로그래밍 언어를 사용하여 일회용 비밀번호(OTP, One-Time Password)를 생성하고 검증하는 라이브러리로, 다양한 2단계 인증(2FA) 시스템에서 사용되는 TOTP와 HOTP 알고리즘을 지원한다. PyOTP는 Google Authenticator, Authy, Microsoft Authenticator와 같은 애플리케이션과 호환되며, 웹 애플리케이션, 모바일 애플리케이션, 서버 애플리케이션에서 보안을 강화하기 위해 자주 사용된다.

2.3 Flask

Flask는 Python으로 작성된 마이크로 웹 프레임워크로 경량화된 디자인을 채택하여, 필요한 기능만을 모듈화된 방식으로 추가할 수 있어 매우 유연하다. 기본적으로 라우

팅, 템플릿 엔진, 개발 서버와 같은 웹 애플리케이션 구축에 필수적인 기능을 제공하며, 확장성을 통해 데이터베이스 통합, 인증, RESTful API 등 다양한 기능을 구현할 수 있다. Flask는 단순하고 직관적인 API를 제공하므로, 빠르게 프로토타입을 개발하거나 소규모 프로젝트에 적합하다.

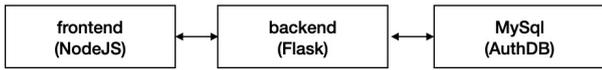
3. 설계

3.1 시스템 구성

관리 시스템은 프론트엔드와 백엔드로 구성되며, 백엔드에서 Flask와 PyOTP를 이용하여 OTP 관리 기능을 수행하고, 프론트엔드인 웹 인터페이스에서 RESTful API를 이용하여 백엔드의 기능을 호출하는 방식으로 구성된다.

TOTP를 이용한 2단계 인증은 비밀키와 현재 시각의 조합으로 인증코드를 생성한다. 사용자에게 비밀키를 발급하고 이를 인증서버와 사용자 단말기에서 공유함으로써 같은 인증코드를 생성할 수 있다.

추가로 OTP를 이용하고자 하는 사용자의 접근 권한을 관리함으로써, 유효한 사용자만 인증받을 수 있도록 설계하였다.



(그림 1) 시스템 구성도

3.2 관리 시스템 요구사항

OTP 인증을 관리자들이 사용자 인증에 활용하기 위해서 기본적으로 요구되는 관리 기능은 다음과 같다.

(표 1) 관리 시스템 요구사항

기능명	내용
비밀키 발급	OTP 비밀키 생성(사용자별 비밀키 발급)
비밀키 관리	사용자별 비밀키 관리(실패 횟수, 유효성 여부 등)
비밀키 폐기	사용자의 비밀키 폐기
인증코드 검사	사용자 아이디와 인증코드의 일치 여부 검사

3.3 RESTful API 구성

OTP 관리를 위한 기능을 Flask를 이용하여 API 서버로 구성함으로써, OTP 관리 시스템뿐만 아니라, 사용자 포털과 같은 다른 웹 서비스를 이용하여 OTP 발급 요청 및 상태 정보 등을 사용자에게 제공할 수 있다.

(표 2) 관리 시스템 RESTful API 정보

URI	기능
/register/{USER}	USER의 비밀키를 생성하고, DB에 저장
/list	등록된 사용자 및 OTP정보 조회
/unregister/{USER}	USER의 OTP 사용 권한 해제
/auth/{USER}/{CODE}	사용자 인증코드 유효성 검사

OTP 비밀키 관리를 위한 정보를 아래 표와 같이 구성하고, 사용자의 OTP 활용 기간 제한 및 실패 횟수 관리 등에 활용할 수 있을 것이다.

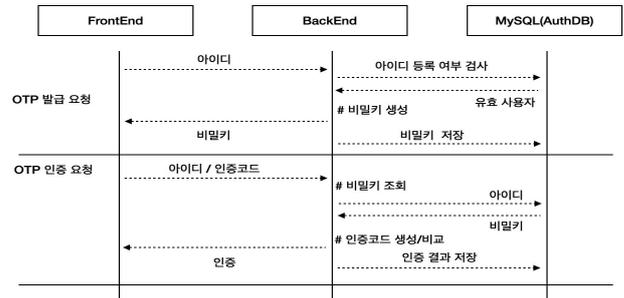
(표 3) 사용자 정보 저장 테이블 정보

컬럼명	내용
seqno	발급 순번(auto_increment)
userid	사용자 아이디
secretkey	OTP 비밀키
issue_date	발급 일시
is_valid	접속 유효성
fail_count	접속 실패 횟수
start_date	OTP 사용 시작일자
end_date	OTP 사용 만료 일자

4. 구현

3.1 API 동작 흐름

아래 그림은 OTP 발급과 인증과정의 API 동작 흐름을 보여주는 예시이다.



(그림 2) 시스템 동작 흐름

아래 그림은 관리 페이지 화면으로, 사용자별 비밀키 정보와 계약 정보 등을 보여주는 화면이다.

사용자 등록 OTP 등록 현황 환경 설정							
No	사용자 아이디	비밀키	발급일시	계약시작일	계약만료일	계약 상태	실패 횟수
1	user01	HIQDESDBAQ4HBJQWKKXLXOKZBGQ6MMYX	2024-08-05 14:54:38	2024-08-01	2024-12-31	Open	0
2	user02	KWPYSQ7WNTG8BAUENM6G7HEUKPNJ662Y	2024-08-10 14:54:38	2024-08-01	2024-11-31	Close	0
3	user03	WD3QU5C2MFUJ3AD2QCXYTKZTPQVOKH7	2024-07-05 14:54:38	2024-07-01	2024-10-31	Open	2
4	user04	HLMQVGR7AYIDH2FJWLJ4LCKGF47LJUY	2024-09-05 14:54:38	2024-09-01	2024-12-31	Open	1
5	user05	UTKVOVS67QUNKXJ557LZ4M54AT7L6I6	2024-09-05 14:54:38	2024-09-01	2024-12-31	Open	0
6	user06	ZX383U6GU5QGPFUVPDDSUOAJFB3PMFCM	2024-09-05 14:54:38	2024-09-01	2024-12-31	Open	0

(그림 3) OTP 관리 페이지

5. 결론

본 고에서 개발한 시스템은 관리 도구의 부재로 인해서 OTP 도입에 어려움을 겪는 시스템 관리자들에게 효율적인 관리 환경을 제공할 것이다. API 방식을 지원함으로써 다양한 웹 서비스와 쉽게 연동을 할 수 있다. 서비스 환경에 따라서, 사용자 관리 및 시스템 접근 권한 설정과 같은 정보를 추가/확장함으로써 다양한 서비스 환경 구축에 이바지할 것이다.

참고 문헌

- [1] 용승림, 최영철, 윤현석, "google OTP를 이용한 스마트 도어락 시스템" 『한국컴퓨터정보학회 하계학술발표논문집』 제24권 제2호, 2016. pp.223-224.

응급환자 이송을 위한 해양원격의료지도시스템 실증 사례 분석

박건우, 최종명, 김선경, 이영호
목포대학교 재난안전사업단, 목포대학교 간호학과, 목포대학교 컴퓨터공학과
e-mail :youngho@mokpo.ac.kr

Case Analysis of a Maritime Telemedical Direction System for Transferring Emergency Patients

Gunwoo Park, Jongmyung Choi, Sun Kyung Kim, Youngho Lee
Mokpo National University

요 약

본 연구는 해양경찰의 응급환자 이송을 위한 원격의료지도시스템을 개발하고, 세 차례의 실증 연구를 통해 시스템의 성능과 효율성을 평가하였다. 첫 번째 실증에서는 음성 통신 문제, 스마트글래스 연결 지연 등의 초기 문제를 발견하고 개선하였다. 두 번째 실증에서는 사용자 정보 부족과 다수 의사의 동시 연결 실패 문제 등을 해결하여 시스템의 안정성과 사용성을 높였다. 마지막 실증에서는 심전도 데이터 표출, 웹 화면 데이터 업데이트 등 고급 기능의 성능을 평가하고, 추가적인 개선을 진행하였다. 연구 결과, 시스템이 해양경찰의 응급환자 이송 작업에 효과적으로 기여할 수 있는 가능성을 제시하며, 향후 추가적인 기능 개선과 장기적인 성능 평가의 필요성을 강조하였다.

1. 서 론

해경의 응급환자 이송은 환자의 생명과 안전을 지키기 위한 매우 중요한 과정이다. 이러한 과정은 지리적 제약과 시간적 압박으로 인해 의료 지원의 적시성과 적절성을 보장하는 데 어려움을 겪을 수 있다[1]. 이러한 문제를 해결하기 위한 방안으로 원격의료지도시스템의 도입이 점차 주목받고 있으며, 특히 스마트글래스와 같은 첨단 기술을 활용한 시스템이 주목받고 있다. 스마트글래스는 실시간으로 영상과 음성을 전달할 수 있어 현장 상황을 정확히 파악하고 적절한 의료 지시를 내리는 데 유용한 도구가 될 수 있다[2].

원격의료지도시스템은 이러한 문제를 극복하고 실질적인 효과를 검증하기 위한 실증 연구가 필요하다[3]. 본 연구는 해경 응급환자 이송을 위한 원격의료지도시스템의 실증을 통해 발견된 문제점과 개선 사항을 분석하는 것을 목표로 하고 있다. 연구는 시스템의 신뢰성과 효율성을 평가하며, 실무에서의 적용 가능성을 높이기 위한 방향을 제시하고자 한다[4].

2. 연구 설계 및 실증 과정

본 연구는 해경의 응급환자 이송을 위한 원격의료지도시스템의 개발과 실증을 목표로 하며, 시스템의 성능과 효율성을 평가하고 개선 방안을 제시하고자 하였다. 연구는 총 3회의 실증으로 구성되었으며, 각 실증 단계에서 시스템의 기능을 점검하고 문제점을 도출하여 개선 사항을 반영하였다.

첫 번째 실증은 2023년 4월 실시되었으며, 초기 시스템의 기능 검토와 문제점 도출을 목표로 하였다. 이 단계에서는 총 20회의 원격의료지도 시도가 이루어졌고, 이 중 8건이 성공적으로 완료되었다. 발견된 주요 문제점으로는 의료지도 요청받는 기관의 전담 인력 부족, 음성 통신 문제, 스마트글래스의 통신 연결 지연, 의료기기 연결의 어려움, 보조 배터리 문제, 음량 조절 필요성이 있었다. 이 문제들을 해결하기 위해 음성 알림 기능을 추가하고, 통신 방법을 LTE방식과 Wifi방식 2개 중 상황에 맞게 접속할 수 있게 개선하였으며, 사용자 매뉴얼을 보강하고, 보조 배터리의 관리 및 조작을 간단하게 할 수 있도록 하였다. 이러한 조치를 통해 시스템의 초기 문제를 최소화하였다.

두 번째 실증은 2023년 6월 진행되었으며, 첫 번째 실증에서 발견된 문제점의 개선과 추가 기능의 평가를 목표로 하였다. 총 41회의 원격의료지도 시도 중 11건이 완료되었고, 이 단계에서의 주요 문제점으로는 의료지도 요청 시 요청자의 정보 부족, 다수 의사의 동시 수락 시 연결 실패, 지도 의사의 화면 미표출, 앱의 강제 종료 문제, 스마트글래스와의 연결 문제 등이 있었다. 이를 해결하기 위해 청각적 알림 기능을 추가하고, 의료지도 요청자의 정보를 표시하여 의료지도 의사의 모습을 볼 수 있게 하였으며, 다수 의사 동시 수락 문제를 해결하기 위해 요청 순서 기반으로 연결하고, 연결되지 않은 의사는 대기할 수 있도록 하였다. 또한, 현장 인력 부족과 응급 상황을 고려하여 스마트폰 없이 즉시 연결 가능한 앱을 개발하였다. 이러한 개선을 통해 시스템의 사용성과 안정성을 높였다.

세 번째 실증은 2023년 7월 진행되었으며, 개선된 시스템의 최종 평가와 실제 사용 상황에서의 성능 검토를 목표로 하였다. 이 단계에서는 총 60건의 원격의료지도를 진행하였으며, 주요 문제점으로는 심전도 데이터 표출 필요성, 웹 화면의 데이터 업데이트 문제, 의료키트 실시간 정보 저장 문제, 스마트글래스의 발열 문제가 발견되었다. 이에 따라 12 유도 심전도 데이터 표출 기능 추가, 웹 데이터 변경 시 화면 자동 업데이트 기능을 개선, 의료키트 실시간 정보 저장 및 일지 표출 기능을 강화, 스마트글래스의 발열 문제 등을 현재 해결 중에 있다.

이러한 일련의 실증 과정과 개선 활동을 통해 원격의료지도시스템의 신뢰성과 효율성을 높이고, 현장 사용에 적합한 시스템으로 발전시키기 위한 기반을 마련하였다. 본 연구는 원격의료지도시스템이 실무에서 효과적으로 적용될 수 있는 방향을 제시하며, 향후 해경의 응급환자 이송 작업에 성공적으로 통합될 수 있는 가능성을 평가하였다.

3. 논의

각 실증 단계에서 발견된 문제점과 그 해결 방안이 시스템의 안정성과 효과성을 어떻게 향상시켰는지 구체적으로 분석하였다. 첫 번째 실증에서의 주요 문제점은 시스템의 초기 설정 및 기능 구현과 관련된 것이었으며, 이 문제를 해결하기 위한 개선 조치들이 시스템의 기본적인 기능성을 확보하는 데 도움을 주었다. 두 번째 실증에서는 사용자 정보 부족과 다수 의사 동시 수락 문제 등의 기능적 한계가 드러났고, 이에 대한 개선 조치는 시스템의 실용성을 높이는 데 기여하였다. 마지막으로, 세 번째 실증에서는 심전도 데이터 표출 및 웹 화면 업데이트 문제 등 고급 기능의 성능이 평가되었으며, 이러한 문제의 해결은 시스템의 전체적인 신뢰성과 유용성을 강화하는 데 중요한 역할을 하였다.

실무 적용 가능성 측면에서, 본 연구의 원격의료지도시스템은 해경의 응급환자 이송 작업에 실질적으로 기여할 수 있는 잠재력을 가지고 있다. 개선된 시스템의 기능들이 현장 상황에서 효과적으로 작동하며, 응급 상황에서의 의료지도를 보다 신속하고 정확하게 제공할 수 있도록 도와줄 것으로 기대된다. 그러나 시스템의 완전한 실무 적용을 위해서는 추가적인 테스트와 지속적인 개선이 필요할 것이다.

4. 결론

본 연구는 해경의 응급환자 이송을 위한 원격의료지도시스템의 개발과 실증을 통해 시스템의 성능을 평가하고, 발견된 문제점을 개선하는 데 기여하였다. 연구 결과, 각 실증 단계에서의 문제점과 개선 조치가 시스템의 신뢰성과 효율성을 높이는 데 중요한 역할을 하였으며, 향후 해경의 응급환자 이송 작업에 성공적으로 통합될 수 있는 가능성을 제시하였다. 본 연구는 원격의료지도시스템의 실무에서의 효과적인 적용을 위한 방향을 제시하며, 실무에

서의 활용 가능성을 평가하였다. 향후 연구에서는 추가적인 기능 개선과 장기적인 성능 평가를 통해 시스템의 완성도를 높이고, 다양한 응급 상황에서도 안정적으로 작동할 수 있는 방안을 모색할 필요가 있다.

ACKNOWLEDGMENTS

본 논문은 행정안전부 지역맞춤형 재난안전 연구개발 사업의 지원을 받아 수행된 연구임(20025162)

참고문헌

- [1] Lee, Y., Kim, S. K., Yoon, H., Choi, J., Go, Y., & Park, G. W. (2022). Smart glasses and telehealth services by professionals in isolated areas in Korea: Acceptability and concerns. *Technology and Health Care : Official Journal of the European Society for Engineering and Medicine*, 1 - 11. <https://doi.org/10.3233/THC-220242>
- [2] Choi, J., Kim, S. K., Lee, Y., Yoon, H., Go, Y., & Byun, K. S. (2021). Perceived Usefulness and Attitude toward Smart-glass for First-aid Remote Support among Coast Guards in Korea. *Journal of the Korea Industrial Information Systems Research*, 26(4), 1-9.
- [3] Yang, H. M., Namgoong, M., & Kim, G. Y. (2022). A Study on the Functionality, Utilization, and Convenience of the Coast Guard Remote Emergency Medical System for Coast Guards. *Journal of the Korean society for Wellness*, 17(4), 165-171. [10.21097/ksw.2022.11.17.4.165](https://doi.org/10.21097/ksw.2022.11.17.4.165)
- [4] Lee, Y., Hwang, I. C., Yang, H. M., Park, G. W., Lee, S. (2024). Design and implementation of a smart glass-based emergency tele-medical direction system. *Smart Media Journal*, 13(5), 26-32. [10.30693/SMJ.2024.13.5.26](https://doi.org/10.30693/SMJ.2024.13.5.26)

실시간 의미 그래픽을 이용한 자동 새우 크기 측정

와카르 말릭 무하마드^{1,2} 김상철^{*,2}

¹전북대학교 전자정보공학부

²전북대학교 지능형로봇연구소

malikwaqarhaider@gmail.com, *sckim7777@gmail.com

Automated Shrimp Size Measurement Using Semantic Graphics in Real Time

Waqar Malik Muhammad^{1,2}, SangCheol Kim²

¹Department of Electronics and Information Engineering, Jeonbuk National University

²Core Research Institute of Intelligent Robots, Jeonbuk National University

Abstract

Accurate measurement of shrimp length is crucial for marine biology and aquaculture, where it is vital for assessing growth, health, and maturity. Traditional manual measurement techniques are often time-consuming and prone to errors, necessitating the development of automated, efficient, and reliable methods. This paper introduces an innovative pipeline for automated shrimp length measurement using advanced computer vision techniques. Our method starts with instance segmentation via Mask R-CNN, which accurately identifies and isolates individual shrimp from their environment. We then apply semantic graphics algorithms to predict the center line, or medial axis, of each shrimp, which forms the geometric basis for length measurement. By calculating the pixel-to-length ratio based on this centerline, we obtain precise length estimates. Experimental results demonstrate that our approach delivers superior precision and consistency compared to both traditional manual methods and existing automated systems.

Keywords: Computer Vision, Semantic Graphics

1. Introduction

Significant advances have been made in developing non-invasive, automated techniques for aquatic species size estimation, particularly using image-based approaches like stereo vision and 3D imaging due to their efficiency and accuracy [1]. With the integration of deep learning and artificial intelligence, these methods have become more precise, allowing for automatic detection and size estimation from underwater images [3]. While much progress has been made for larger species, such as fish, less attention has been given to smaller species like shrimp, which are crucial in aquaculture.

Automated and accurate measurement of aquatic organisms is essential for effective stock management and growth monitoring. Traditional manual methods, though precise, are labor-intensive, time-consuming, and invasive [1]. As a result, non-invasive techniques such as 2D and 3D imaging have gained prominence, leveraging advancements in computer vision for fast and accurate measurements. For example, stereo vision has been widely used for fish size estimation [2]. The application of deep learning has further enhanced these methods, with convolutional neural networks (CNNs) proving effective in automating size estimation from underwater images [3]. Other technologies like sonar and LiDAR have also been explored for volumetric analysis and size estimation [4], contributing to more efficient aquaculture management.

This paper introduces a novel pipeline for automated shrimp length measurement using advanced computer vision techniques. Our approach uses Mask R-CNN for instance segmentation to detect individual shrimp and applies semantic graphics algorithms to predict the medial axis, or center line, for accurate length calculation. Experimental results show that our method surpasses traditional and existing automated systems, offering higher precision and consistency, which is critical for shrimp growth and health assessment in aquaculture.

2. Materials and Methods

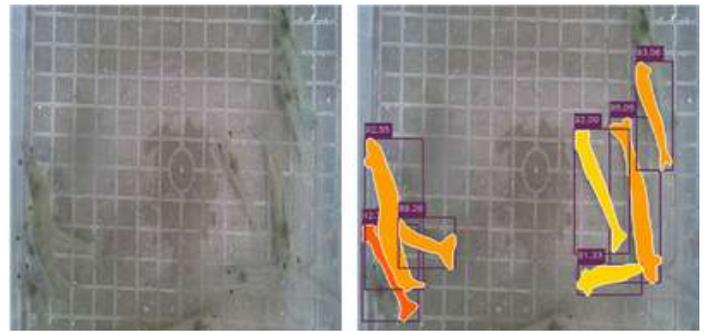
To develop a robust model for shrimp length estimation, we collected a comprehensive dataset of shrimp images from four distinct environments. The first two environments contained 6 and 12 shrimps, respectively, while the third and fourth environments housed between 12 and 24 shrimps. Lighting conditions varied across environments: the first two offered favorable lighting, while the latter two were slightly darker, simulating real-world variability.

For data collection, a plastic container measuring 50 cm in length and 35 cm in width was used, with a camera mounted directly above the container. The camera, operating at 30 frames per second (FPS), captured high-resolution images of the shrimp. An initial dataset of 1,000 images was obtained, covering a range of shrimp counts and lighting conditions. Each image was

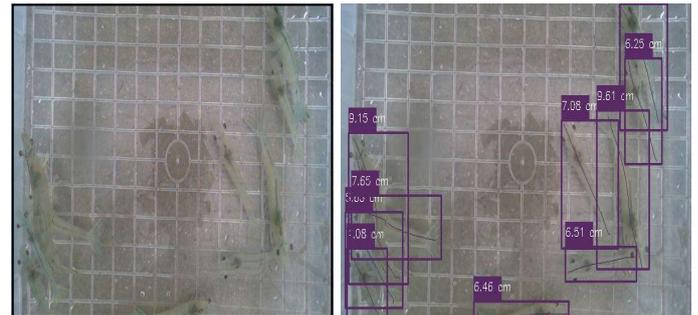
meticulously labeled to enable accurate instance segmentation of individual shrimp for training the deep learning models.

To detect and isolate individual shrimps, we trained an instance segmentation model using Mask R-CNN. The model was optimized on the labeled shrimp dataset, with performance evaluated using the bounding box mean average precision (Box mAP). The model achieved a high accuracy score of 0.966, demonstrating its ability to accurately detect shrimp across varying lighting conditions and shrimp densities as shown in figure 1. Following instance segmentation, a image segmentation model was applied to the masks generated by Mask R-CNN. This model was trained to extract the center line of each shrimp from the segmentation masks. Given the known height of the camera, we computed the length of each shrimp's center line using pixel-to-length ratios.

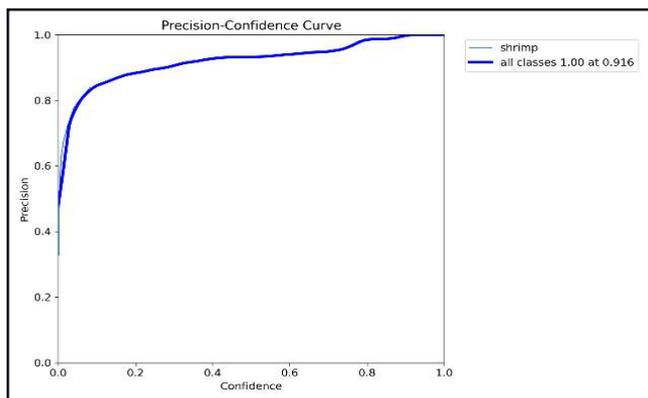
The results of the shrimp length estimation, along with performance metrics and accuracy, are presented in the following sections. These findings highlight the effectiveness of combining instance and semantic segmentation for precise shrimp length measurement. The flow diagram for the prediction of shrimp length is presented in the figure 2. The results of Instance segmentation model are provided in the figure 3.



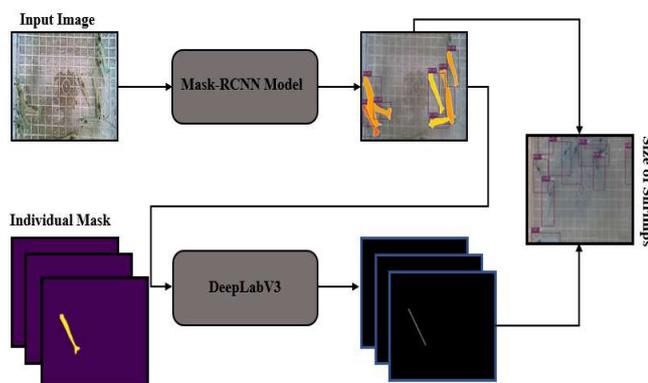
(Figure 3) Results of Instance Segmentation Model



(Figure 4) Results of Instance Segmentation Model



(Figure 1) Training precision-confidence curve of Mask-RCNN



(Figure 2) Flow Diagram for the prediction of shrimp length

The figure 4 is showing the length prediction of individual shrimps with bounding boxes.

Acknowledgment

This research was supported by the Regional Innovation Strategy (RIS) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2023RIS-008), the Agricultural Policy Support Technology Development Project funded by the Rural Development Administration's National Institute of Crop Science (RS-2023-00232338), and the University-Focused Research Institute Support Program funded by the National Research Foundation of Korea (NRF-2019R1A6A1A09031717).

References

- [1] Espinosa-Faller, F. J., García, M. A., & Rendón-Mancha, J. M. (2017). Non-invasive fish size estimation using stereoscopic vision. *Journal of Applied Aquaculture*, *9*(9), 4795-4807. <https://doi.org/10.1111/are.13322>
- [2] Costa, C., Antonucci, F., Pallottino, F., & Menesatti, P. (2011). Fish morphometric measurement from digital images using an automatic software. *Journal of Applied Aquaculture*, *2*(2), 45-50. <https://doi.org/10.1016/j.aquaeng.2010.11.003>
- [3] Liu, Z., Huang, J., & Su, W. (2019). Automatic fish size estimation from underwater images using deep learning. *Journal of Applied Aquaculture*, *5*(5), 1178-1186. <https://doi.org/10.1007/s00773-019-00653>
- [4] Ditrach, T., Jirina, M., & Jirina, M. Jr. (2015). Object detection and fish size estimation for aquaculture using image processing. *Journal of Applied Aquaculture*, *4*(4), 351-356. <https://doi.org/10.1016/j.ifacol.2015.07.063>

검색 증강 생성의 효과성에 대한 탐구

Athiruj Poositaporn^{1,2}, 정한민^{1,2}, 박정훈¹
¹한국과학기술정보연구원 융합서비스센터
²과학기술연합대학원대학교 응용AI 전공
 {athiruj, jhm, jhpark}@kisti.re.kr

Study on the Effectiveness of Retrieval-Augmented Generation

Athiruj Poositaporn^{1,2}, Hanmin Jung^{1,2}, Jung Hoon Park¹
¹Convergence Service Center, Korea Institute of Science and Technology Information
²Applied AI Major, University of Science and Technology

Abstract

Advancements in Natural Language Processing (NLP) and Large Language Models (LLMs) have revolutionized text generation and enhanced the capability to produce more accurate and contextually relevant content. This study examines the ability of Retrieval-Augmented Generation (RAG) by manually reviewing the generated answers against a basic GPT-4 model. In the experiment, we generated RAG results on different chunk sizes (K = 5, 10, and 20) and compared them to a result from GPT-4. The results indicated that RAG provided more detailed and contextually enriched responses compared to GPT-4. In future works, we aim to refine indexing, query translation, and retrieval techniques and expand the application across different datasets and domains.

1. Introduction

Recent advancements in Natural Language Processing (NLP) have resulted in significant improvements in text generation tasks[1]. This is due to the development of Large Language Models(LLMs) such as GPT-4 and Retrieval-Augmented Generation(RAG) models[2-3]. While GPT-4 generates human-like text solely through extensive pre-training on diverse datasets, RAG models improve text generation by incorporating retrieval mechanisms that extract relevant information from external knowledge bases[4]. In this study, we focus on exploring the effectiveness of RAG models by investigating their ability to generate contextually relevant answers by using external knowledge. We also evaluate against a basic GPT-4 as a benchmark. Therefore, we aim to investigate the generated outputs through manual evaluation to assess how well RAG integrates retrieved information along with a predefined prompt to enhance answer quality.

2. Design

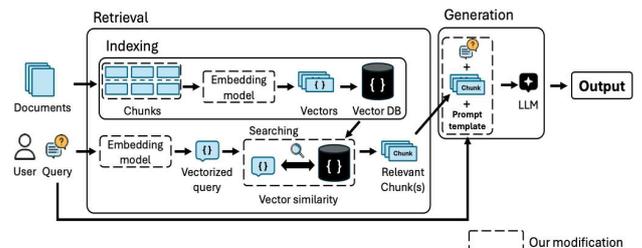
In this study, we aim to compare the performance of a RAG model with a basic GPT-4 model in generating contextually relevant answers. The specific details and configuration of the RAG model used in this study are presented in the following section.

2.1 Our RAG Model

We based our RAG model and made several modifications to

the original paper[5]. Figure 1 shows the two critical processes of our RAG model: retrieval and answer generation. The retrieval process extracts relevant information from external knowledge sources at the sentence level by using Natural Language Toolkit(NLTK)¹⁾ to create chunks of three sentences. These chunks are embedded using the text-embedding-3-large model²⁾ with a set number of chunks (K = 5, 10, and 20). The retrieval process then evaluates their relevance to the embedded query through cosine similarity.

The answer generation process involves combining the retrieved information into a predefined prompt template³⁾(See Figure 2) and using OpenAI's GPT-4 model to generate a coherent response to the query. This process is significantly influenced by how the predefined prompt template is created to guide the format and content of the answers. We configure the GPT-4 model with a temperature parameter set to 0 to minimize randomness and ensure consistent output.



(Figure 1) Our RAG Model

¹⁾<https://www.nltk.org/>

²⁾<https://platform.openai.com/docs/guides/embeddings/>

³⁾<https://community.openai.com/t/prompt-engineering-for-rag/621495/2>

2.2 Dataset

The dataset for this study is from the book "AI for Climate Change and Environmental Sustainability," published in PDF format on August 14th, 2024, containing 179 pages. Notably, the external data used for retrieval includes information that is unseen to the GPT-4⁴⁾ in the generation process, allowing us to determine the performance of RAG in generating answers.

3. Discussion

In our investigation, we applied a predefined prompt template to ensure consistency across different retrieval settings. We found that the higher chunk sizes (K = 10 and 20) of RAG provide more detailed and contextually enriched responses compared to GPT-4, as an example of the comparison shown in Table 1. It is noteworthy that even when the retrieved data included information not previously seen by GPT-4, the RAG model was still able to generate coherent, contextually relevant answers, demonstrating its ability to effectively utilize external knowledge.

4. Conclusions

This study explored the effectiveness of Retrieval-Augmented Generation (RAG) models in generating contextually relevant answers by incorporating external knowledge.

We found that the RAG model improved over GPT-4 in providing more detailed and enriched responses. The design of the predefined prompt was crucial in guiding the model to generate coherent and accurate content, demonstrating its importance in text generation quality.

Additionally, the model effectively generated coherent and enriched content even with unseen information by GPT-4, enhancing its text generation capabilities.

In future work, we aim to explore the potential of fine-tuning the RAG model to further improve its performance by investigating query translation, retrieval optimization, and indexing strategies. Additionally, we plan to conduct more extensive evaluations to assess the RAG model's capabilities across a wider range of datasets and domains.

Acknowledgement

* 본 연구는 2024년도 수행된 KISTI 주요사업인 '지능형 과학기술지식인프라 통합서비스 플랫폼 구축 (K24L1M3C4-01)'을 통해 수행되었습니다.

References

- [1] M. Bayer et al., "Data augmentation in natural language processing: A novel text generation approach for long and short text classifiers," International journal of machine learning and cybernetics, vol. 14, no. 1, pp. 135 - 150, 2023.
- [2] PP. Ray, "ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope," Internet of Things and Cyber-Physical Systems, vol. 3, pp. 121 - 154, 2023.
- [3] J. Chen, H. Lin, X. Han, and L. Sun, "Benchmarking large language models in retrieval-augmented generation," In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, no. 16, pp. 17754 - 17762, 2024.
- [4] P. Zhao et al., "Retrieval-augmented generation for ai-generated content: A survey," arXiv:2402.19473, 2023.
- [5] P. Lewis et al., "Retrieval-augmented generation for knowledge intensive NLP tasks," In Proceedings of 34th Advances in Neural Information Processing Systems, pp. 9459-9474, 2020.

<p>QUESTION: {QUESTION}</p> <p>DOCUMENT: {DOCUMENT}</p> <p>INSTRUCTIONS: You are an AI language model designed to assist with generating informative and accurate content. Answer the user's QUESTION using the information provided in the DOCUMENT. Ensure that your response is factually accurate and directly references the content of the DOCUMENT. Do not include any information that is not present in the DOCUMENT.</p> <p>GUIDELINES: - Start with a concise answer. - Provide evidence/additional information from the DOCUMENT to support your answer. Use bullet points for clarity only if needed. - If the DOCUMENT does not contain sufficient information, clearly state that.</p> <p>REMEMBER: - Use only the information from the DOCUMENT to construct your answer. - Be concise and to the point.</p>
--

(Figure 2) Predefined Prompt Template

(Table 1) Example of a Generated Answer from RAG (K = 20) and GPT-4

Query: How can AI-based tools and technologies help in reducing energy emissions and removing CO2 from the atmosphere?	
RAG (K = 20)	GPT-4
Improving Carbon Capture and Removal: AI can enhance the design and operation of Direct Air Capture (DAC) technologies and Carbon Capture and Storage (CCS) systems, thereby speeding up the development of carbon-neutral solutions.	Carbon Capture: AI can help in designing and optimizing processes for carbon capture and storage. It can also help in monitoring and verifying the effectiveness of these processes.

⁴⁾<https://platform.openai.com/docs/models/gpt-4-turbo-and-gpt-4>

과학기술 이슈 대응 정보서비스의 이용 현황 분석

윤정선, 배수영, 한상준, 유수현
 한국과학기술정보연구원 융합서비스센터
 jsyoon@kisti.re.kr, sybae@kisti.re.kr, sjhan@kisti.re.kr, yoosu@kisti.re.kr

Usage Analysis of a Science and Technology Issue-Responsive Information Service

Jungsun Yoon, Suyeong Bae, Sangjun Han, Suhyeon Yoo
 Korea Institute of Science & Technology Information, Convergence Service Center

요약

매일같이 쏟아져 나오는 정보의 홍수 속에서 새롭게 떠오르는 이슈를 신속하게 파악하고 새로운 지식을 습득하는 것은 과학기술 분야 연구개발을 위해 필수적인 일이다. 본 연구에서는 최신 이슈에 대한 다양한 정보를 큐레이션하여 제공하는 과학기술 이슈대응 서비스 운영 데이터를 기반으로 이용자들이 관심을 많이 갖는 이슈들이 어떤 것인지 분석해 보았다. 2021년 12월부터 2024년 8월까지 생성된 107개의 이슈에 대해 분석한 결과 산불, 블록체인, 황사, 플라스틱 재활용 순으로 조회가 많이 이루어졌다. 이슈의 분야별 분석 결과 보건 의료, 지구과학, 생명과학 순으로 많은 이슈가 생성되었으며, 조회수는 지구과학, 정보통신, 생명과학 순으로 높게 나타났다. 각 이슈 콘텐츠 세부 항목의 이용도 분석 결과 웹뉴스, 대표저널/프로시딩, 키워드 정의, 논문 순으로 클릭수가 높게 나타났다. 본 연구 결과를 통해 이용자들이 어떤 이슈에 관심을 많이 보이는지, 어떤 콘텐츠 항목을 많이 이용하는지 파악할 수 있었다. 본 결과는 시의성 있는 이슈대응 정보서비스의 개발 및 운영에 도움이 되리라 기대된다.

1. 서론

급변하는 세상 속에서 새롭게 부상하는 이슈들을 파악하고 지식을 습득하여 활용하는 것은 필수적이며 중요한 일상이 되어가고 있다. 새로운 이슈에 대한 인사이트를 쉽게 발견할 수 있도록 많은 정보를 종합 분석하여 정보서비스가 제공된다면 연구개발 활동의 효율성을 크게 높일 수 있을 것이다.

본 연구에서는 과학기술 이슈대응 정보서비스의 이용현황 분석을 통해 이용자들이 관심을 많이 가지는 이슈는 어떤 것들이 있으며, 어떤 종류의 콘텐츠를 선호하는지 분석해 보고자 한다. 이슈대응 정보서비스의 이용 분석은 이용자 타게팅 서비스 개발에 참고가 될 수 있을 것이다.

2. 연구 내용 및 방법

2.1 연구 대상

과학기술분야 이슈대응 정보서비스로는 한국지식재산연구원의 주간 지식재산 동향뉴스[1], 한국과학기술기획평가원이 운영하는 글로벌 과학기술정책정보서비스(S&T GPS)[2], 한국과학기술정보연구원의 ScienceON 트렌드 서비스가 있다[3]. 본 연구에서는 해당 이슈를 전반적으로 이해하기 위해 다양한 정보원으로부터 종합적으로 정보를 패키징하여 제공하는 서비스인 ScienceON 트렌드 서비스를 대상으로 연구를 수행하였다. 과학기술 지식인프라 통합서비스인 ScienceON은 다양한 지식인프라를 활용하여

다각도의 정보를 확보하고, 연계융합 기술로 데이터 및 분석정보를 제공하여 이슈에 대한 심도있는 정보를 제공하고 있다[4]. 트렌드에 포함된 정보는 키워드, Overview, 웹뉴스, 학술정보, 분석정보, Data/API, 연계정보이며 상세한 내용은 표1과 같다. 각 이슈에 대해 상기한 정보를 카테고리별로 제공하여 이용자가 해당 이슈를 다각도로 이해하고 정보를 활용할 수 있게 하였다.

(표 1) ScienceON 트렌드 콘텐츠 구성

카테고리	상세 내용
키워드	키워드 정의, 워드 클라우드
Overview	관심도 변화, 요약정보, 주요 출원인, 연구자, 기관/단체 홈페이지, 대표 저널/프로시딩
웹뉴스	이슈 키워드와 연관된 웹뉴스
오피니언	웹뉴스로부터 추출된 전문가 발언
학술정보	논문, 특허, 보고서, 동향
분석정보	COMPAS 핵심경쟁자 탐색과 특허분석, TOD 제품분석
Data/API	DataON, 공공데이터포털의 Data와 API
연계정보	외부 동향자료, 동영상

2.2 이슈대응 정보서비스 운영 현황

트렌드 서비스는 2021년 12월부터 시작하여 2024년 8월까지 104개의 트렌드 콘텐츠를 생성하였다. 처음에는 비정기적으로 생성하다가 2024년 3월부터는 주1회 정기적으로 트렌드 콘텐츠를 생성하고 있다. 수집한 정보의 키워드 분

석을 통해 키워드 후보가 자동 생성되고, 운영자는 그 중 콘텐츠를 생산할 이슈 키워드를 선정한다. 이슈 키워드를 선정한 후 자동화 프로세스에 의해 트렌드 콘텐츠가 생성되면 운영자가 각 내용을 확인하고 적합성을 판단하여 정보를 선별한다. 중복된 정보는 제거하고 외부에서 유입된 정보 중 유료 정보, 상업적 정보도 제외한다. 자동 생성된 정보가 미흡한 경우 이슈에 적합한 정보를 검색하여 교체하고 썸네일 이미지와 동영상 콘텐츠를 선택하는 등의 작업이 수행된다. 기본적인 프로세스는 상당 부분 자동화하였지만, 이슈 키워드의 선정, 상세 콘텐츠 선별과 보완 과정에 운영자에 의한 정보 큐레이션이 이루어지고 있다.

2.3 분석 방법

과학기술 이슈 대응 정보서비스 분석을 위해 트렌드 운영 데이터를 기반으로 많이 이용되는 세부항목, 관심도 높은 이슈 키워드, 분야별 이슈 선정 빈도를 분석하였다.

3. 연구 결과

트렌드 서비스는 2021년 12월부터 서비스가 개시되었으나 개설 후 서브 메뉴 형태로 서비스 되다가 2023년 11월부터 메인페이지에 서비스를 노출되기 시작했다. 그 후 트렌드 서비스의 이용도가 크게 증가하였다.

3.1 세부 항목 이용도 분석

트렌드 상세화면에서 제공되는 항목별 상세 정보를 조회한 클릭수는 2023년 11월 23일부터 수집하고 있다. 상세 정보 클릭수를 분석한 결과 웹뉴스, 대표 저널/프로시딩 순으로 가장 많은 조회가 이루어졌고, 키워드 정의, 논문, 외부, 기관/단체 홈페이지가 그 뒤를 이었다(그림 1) 최신 뉴스를 파악하여 연구를 수행하고 논문을 작성하기 위한 목적으로 트렌드 서비스를 많이 이용하는 것으로 보인다.



(그림 1) 세부 항목별 클릭 수

3.2 관심도가 높은 이슈 키워드

이슈별 조회수 분석 결과 이용자들이 가장 많이 조회한 이슈 키워드는 산불, 블록체인, 황사, 플라스틱 재활용 순이었다. ‘산불’의 경우 2023년 키워드임에도 2024년에 훨씬 많은 조회수가 이루어졌다. 재생에너지, 로봇, 스마트시티도 생성된 연도 이후에 더 많이 조회되었다. 트렌드 콘텐

츠는 메인페이지뿐 아니라 논문, 특허, 보고서 등의 상세 화면에서 키워드가 매칭되는 경우 추천 정보로 제공되고 있다. 트렌드 콘텐츠가 제공 시점 이후에도 지속적으로 이용되고 있음을 확인할 수 있었다.

3.3 분야별 이슈 선정 빈도

선정된 이슈 키워드를 과학기술표준분류에 매핑하여 분석한 결과 그림2와 같이 보건의료, 지구과학, 생명과학, 정보통신 순으로 트렌드 정보가 많이 생성되었다. 주로 뉴스 분석을 통해 시의성 있는 이슈 키워드를 생성하다보니 실생활에 관련 있는 주제들이 많이 선정된 것으로 보인다.



(그림 2) 분야별 이슈 선정 빈도

4. 결론

본 연구에서는 최신 이슈에 대한 다양한 정보를 큐레이션하여 제공하는 과학기술 이슈 대응 서비스 운영 데이터를 기반으로 이용자들이 관심을 많이 갖는 이슈들이 어떤 것인지 분석해 보았다. 2021년 12월부터 2024년 8월까지 생성된 107개의 이슈에 대해 분석한 결과 산불, 블록체인, 황사, 플라스틱 재활용 순으로 조회가 많이 이루어졌다. 이슈의 분야별 분석 결과 보건의료, 지구과학, 생명과학 순으로 많은 이슈가 생성되었으며, 조회수는 지구과학, 정보통신, 생명과학 순으로 높게 나타났다. 각 이슈 콘텐츠 세부 항목의 이용도 분석 결과 웹뉴스, 대표저널/프로시딩, 키워드 정의, 논문 순으로 클릭수가 높게 나타났다. 본 연구 결과를 통해 이용자들이 어떤 이슈에 관심을 많이 보이는지, 어떤 콘텐츠 항목을 많이 이용하는지 파악할 수 있었다. 본 결과는 시의성 있는 이슈 대응 정보서비스의 개발 및 운영에 도움이 되리라 기대된다.

* 본 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다(과제번호: (KISTI) K24L1M3C4-01).

참고 문헌

[1] KIIP. Homepage <https://www.kiip.re.kr>.
 [2] KISTEP. S&T GPS <https://now.k2base.re.kr>.
 [3] KISTI. ScienceON <https://scienceon.kisti.re.kr>.
 [4] 유수현, 배수영. “지식인프라 연계를 통한 이슈 대응 서비스 개발”, 『디지털콘텐츠학회 논문집』 제25권 제5호, 2024. pp.1253-1262.

효율적 조류 객체 탐지를 위한 YOLO 모델 비교 연구

왕흥요¹, 이승재¹, 유지형¹, 유경호², 김형주², 김관구¹
조선대학교 컴퓨터공학과¹, 조선대학교 AI융합연구원²
e-mail : 791157327nn@gmail.com, dl4786@chosun.kr, fbwlgud51@gmail.com,
infinitegh@chosun.ac.kr, hyoungjukim@chosun.ac.kr, pkkim@chosun.ac.kr

Research on comparing YOLO models for efficient avian object detection

WANG XINYAO¹, Seungjae Lee¹, Jihyung Yu¹, Kyungho Yu²,
Hyoungju Kim², Pankoo Kim¹
Department of Computer Engineering, Chosun University¹,
Institute of AI Convergence, Chosun University²

요 약

본 연구는 새꼬막 양식장에서 조류 피해를 예방하기 위한 효율적인 조류 객체 탐지 모델을 비교하였다. YOLOv8 Large와 YOLOv8 Small 모델을 사용하여 실험한 결과, YOLOv8 Large는 일부 이미지에서 더 많은 객체를 정확히 탐지하는 반면, 다른 이미지에서는 YOLOv8 Small이 더 우수한 성능을 보였다. 향후 두 모델의 성능차이를 격차를 줄이는 연구를 진행하여 새꼬막 양식장에서 조류 피해를 최소화 할 수 있는 조류 객체 모델을 선택하고자한다.

1. 서 론

최근 인구가 지속적으로 감소하고있는 농어촌에서 인공지능을 도입하여 사람의 빈자리를 채우는 사례가 증가하고 있다. 전남 고흥 선정마을 에서는 새꼬막 양식사업에 조류피해가 지속되고 있어 피해를 예방하고 주민들의 수입 증대를 위해 인공지능 기술을 활용한 'AI기반 새꼬막 양식어장 관리서비스'가 서비스 되고 있으며 효율적인 조류 객체탐지 방법에 대한 연구가 선행되었다[1].

본 연구에서는 선행연구를 바탕으로 조류 데이터를 YOLOv8 Large, Small모델에 각각 학습하여 두 모델의 차이를 비교하고 효율적인 조류 객체 탐지 모델에 대해 논한다.

2. 관련 연구

2.1 딥러닝 기반 객체 탐지

딥러닝 기반의 객체 탐지 모델은 대표적으로 YOLO, EfficientDet, 그리고 DetectoRS가 있다[2-4]. YOLO는 실시간 객체 탐지에서 높은 정확도와 성능을 자랑하며, EfficientDet은 네트워크의 계산량을 최적화하여 빠른 성능을 제공한다. Google에서 발표한 DetectoRS는 복잡한 객체 탐지에서 비교적 높은 성능을 보이며, 특히 다양한 크기와 형태의 객체를 정확하게 인식할 수 있다. 이러한 딥러닝 기반의 객체 탐지 기술은 자율주행 차량, 얼굴 인식 시스템, 그리고 스마트팜과 같은 다양한 응용 분야에서 광범위하게 사용되고 있다. 스마트팜에서는 식물 상태 모

니터링 및 수확 시기 예측과 같은 작업에서 객체 탐지 기술이 필수적이다. 또한 본 연구와 같이 농사나 양식에서 피해를 예방하는 용도로 사용되고 있다.

3. 딥러닝 기반 YOLOv8 조류 객체 탐지 모델

3.1 YOLOv8 Large, Small

YOLO 모델에서 가장 최신버전은 v8버전이다. YOLOv8은 객체 탐지 분야에서 성능과 효율성을 고려하여 다양한 모델 크기를 제공한다. YOLOv8 Large는 큰 네트워크 아키텍처를 채택하여 높은 탐지 정확도를 나타낸다. 이 모델은 많은 층과 파라미터를 포함하고 있어 복잡한 장면에서도 뛰어난 성능을 발휘하며, 고해상도 이미지에서의 세밀한 객체 인식이 가능하다. 그러나, 이러한 큰 모델은 더 많은 계산 자원과 메모리를 요구하기 때문에 처리 속도가 느리다는 단점을 지니고 있다.

YOLOv8 Small은 네트워크의 크기를 줄여 효율성을 높인 모델이다. 적은 수의 층과 파라미터를 사용하여 계산 자원과 메모리 사용을 최소화하며, 더 빠른 속도로 객체 탐지를 수행할 수 있다. 이 모델은 리소스가 제한된 환경이나 실시간 처리 요구가 있는 애플리케이션에서 유리하다. 하지만, 모델의 경량화로 인해 YOLOv8 Large보다 정확도는 다소 낮을 수 있다. 이처럼 YOLOv8 Large와 YOLOv8 Small은 각각의 장점과 단점을 가지고 있어, 사용하고자하는 분야에 맞는 데이터를 직접 학습시키고 결과를 확인하여 사용하여야 한다. 따라서, 본 연구에서는 새꼬막 양식에서 조류의 피해를 예방하기위해 조류 데이

표 1 . YOLO v8 Large와 Small 모델을 이용한 조류 객체 탐지의 비교결과

구분	(a)	(b)	(c)
YOLO v8 Large			
YOLO v8 Small			

터를 학습하여 두 모델사이의 성능을 비교한다.

3.2 실험 결과

실험결과 그림 1에서 확인 할 수 있듯이 학습횟수에서 정확도는 YOLOv8 Large 모델이 높은 것을 확인할 수 있었다. 하지만 같은 이미지를 YOLOv8 Large, Small 모델에 입력하여 추론한 결과 표 1-(a)에서 볼 수 있듯이 Large 모델이 성능은 뛰어나지만 새가 탐지 되지 않았지만 Small 모델에서는 탐지가 된 것을 확인할 수 있었다.

상대적으로 더 좋은 성능을 보였다. 객체 인식의 정확도 측면에서, 특히 배경과 객체의 색상이 유사할 때 오탐지 또는 미탐지가 발생하는 경우가 있었다. 모델 성능의 효율성 측면에서 YOLOv8 Large는 높은 인식률을 제공하지만, 상대적으로 계산 비용과 시간이 많이 소요된다. 반면, YOLOv8 Small은 상대적으로 빠른 처리 속도를 자랑하지만 인식률이 낮아지는 경향을 보였다. 향후 두 모델의 성능격차를 극복하기 위한 연구를 진행하도록 하겠다.

감사의 글

본과제(결과물)는교육부와한국연구재단의재원으로지원을받아수행된3단계산학협력선도대학육성사업(LINK 3.0)의연구결과입니다. 또한, 본 연구성과물은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. RS-2024-00463238)

참고 문헌

- [1] 유경호, 방민욱, 강건, 송민주, 김관구, “딥러닝기반 드론예찰 시스템에서의 조류객체탐지 연구”, 2023종합학술대회, 스마트미디어학회. 2023
- [2] Terven, Juan, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas." Machine Learning and Knowledge Extraction 5.4 (2023): 1680-1716.
- [3] Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10781-10790).
- [4] Qiao, S., Chen, L. C., & Yuille, A. (2021). Detectors: Detecting objects with recursive feature pyramid and switchable atrous convolution. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10213-10224).

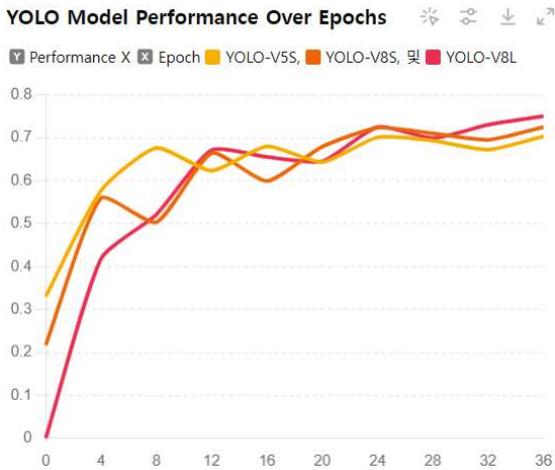


그림 1. YOLOv8 Large와 Small의 학습에 따른 정확도 비교

4. Conclusion

본 연구에서는 새꼬막 양식장에서 조류피해를 예방하기 위한 효율적 조류 객체탐지 모델을 비교하였다. 실험결과 YOLOv8 Large와 YOLOv8 Small 모델 간의 인식률에는 상당한 편차가 나타났다. 동일한 이미지를 처리한 결과, YOLOv8 Large는 일부 이미지에서 더 많은 객체를 정확히 인식하는 반면, 다른 이미지에서는 YOLOv8 Small이

MLCA 기법을 적용한 향상된 조류 객체 탐지모델 연구

왕흥요¹, 이승재¹, 유지형¹, 유경호², 김형주², 김판구¹
 조선대학교 컴퓨터공학과¹, 조선대학교 AI융합연구원²
 e-mail : 791157327nn@gmail.com, dl4786@chosun.kr, fbwlgud51@gmail.com,
 infinitegh@chosun.ac.kr, hyoungjukim@chosun.ac.kr, pkkim@chosun.ac.kr

Research on Enhanced Bird Object Detection Model Using MLCA Technique

WANG XINYAO¹, Seungjae Lee¹, Jihyung Yu¹, Kyungho Yu²,
 Hyoungju Kim², Pankoo Kim¹
 Department of Computer Engineering, Chosun University¹,
 Institute of AI Convergence, Chosun University²

요약

본 연구에서는 풍력 터빈과 새들 간의 충돌 문제를 해결하기 위해 YOLOv8 알고리즘을 기반으로 한 경량 조류 탐지 모델을 제안한다. 기존 YOLO 모델에 MLCA 메커니즘을 적용해 작은 객체 탐지 능력을 향상시켰으며, 실험 결과 모델의 정확도가 1%~1.5% 증가하고 추론 복잡도는 감소하였다. 앞으로는 특징 융합 기술과 새로운 탐지 헤드 기법을 통해 성능을 더욱 향상시킬 계획이다.

1. 서론

최근 전 세계 풍력 발전 기술의 급속한 발전은 풍력 터빈과 새들 간의 충돌 문제에 대한 인식을 높이고 있다. 이러한 충돌은 풍력 발전 장비 유지보수 비용의 상당한 증가를 초래할 뿐만 아니라 조류 생태계에 해로운 영향을 미친다. 기존 연구 데이터에 따르면, 풍력 터빈 주변에서 자주 발견되는 새들, 예를 들어 참새, 갈매기, 오리 등이 치명적인 충돌을 경험한 것으로 나타났다.

이 문제를 해결하기 위해 본 연구는 YOLOv8 알고리즘을 기반으로 한 경량 조류 탐지 모델의 구현을 제안한다. 이 모델은 기존 YOLO 아키텍처의 복잡성을 최소화하여 탐지 정확도를 향상시키고, 따라서 새를 탐지하는 능력을 증가시킨다. 시스템은 드론을 활용하여 조류 활동을 실시간으로 모니터링하며, 새가 풍력 터빈에 접근하면 드론이 자동으로 새의 경로로 비행하여 점진적인 경고음을 발산하여 새가 풍력 터빈에 접근하지 않도록 효과적으로 방지한다. 또한, 이 시스템은 동일한 새의 반복 탐지를 방지하는 기술을 포함하고 있어 드론의 운영 효율성을 향상시키고 인근 주민에 대한 방해를 줄인다.

이 기술의 개발은 풍력 발전 장비를 효과적으로 보호할 뿐만 아니라 새와 그 서식지에 대한 부정적인 영향을 크게 줄이는 데 기여할 것이다. 따라서 본 연구에서는 기존 YOLO 모델을 수정 및 개선하여 조류 개체 탐지를 효율적으로 할 수 있도록 하는 것이 목적이다.

2. 관련 연구

2.1 딥러닝 기반 객체 탐지

최근 몇 년간 객체 인식 기술이 크게 발전하였으며, Faster R-CNN과 DETR과 같은 주목할 만한 모델들이 소개되었다. 그러나 YOLO 모델은 뛰어난 전역 추론 속도와 효율적인 학습 성능 덕분에 이미지 처리 분야에서 여전히 가장 인기 있는 모델 중 하나이다[1]. 최근 YOLO 모델은 버전 8.2로 주요 업데이트가 이루어졌으며, 이는 더 경량화된 아키텍처와 자동화된 파라미터 조정 기능을 도입하였다.

그럼에도 불구하고, 평균 새의 길이가 약 6센티미터인 점을 고려할 때, 현재의 YOLOv8 모델은 큰 객체에 비해 작은 객체(예: 새)를 인식하는 정확도가 여전히 낮다. 따라서 본 연구는 YOLOv8nano 모델을 기반으로 한 새로운 주의(Attention) 메커니즘을 도입하여 YOLO 모델의 작은 객체(새 등) 인식 능력을 향상시키는 것을 목적으로 한다.

3. 혼합 지역 채널 어텐션 메커니즘을 이용한 YOLO모델 개선

3.1 MLCA 메커니즘

MLCA(Mixed Local Channel Attention) 메커니즘은 입력 피쳐 맵을 최적화하기 위해 개발된 것이다[2]. 이 메커니즘은 로컬 평균 풀링(Local Average Pooling, LAP)과 글로벌 평균 풀링(Global Average Pooling, GAP)을 통해 각각 로컬과 글로벌 특징을 추출한다. 이후, 이 특징들은 1D 합성곱(1D Convolution)을 통해 변환되며, 채널 차원은 압축되고 공간 차원은 유지된다. 그 다음, 로컬 풀링된 특징들은 재구성되어 원본 입력 특징들과 곱셈 연산을 통해 결합되며, 유용한 특징에 대한 집중도를 향상시킨다. 글로벌 풀링된 특징들도 1D 합성곱 및 재구성을 거쳐 로

컬 풀링된 특징들과 덧셈 연산을 통해 융합되며, 글로벌 컨텍스트 정보를 통합한다. 마지막으로, 처리된 피쳐 맵은 언스폴링(unspooling) 연산을 통해 원래의 공간 차원으로 복원된다.

3.2 MLCA를 적용한 YOLOv8n

YOLOv8n 모델은 주로 피쳐 융합 레이어와 피쳐 추출 레이어로 구성된다. 우리는 MLCA 주의(Attention) 메커니즘을 백본(피쳐 추출 레이어)에 통합하였으며, 그 결과 새로운 YOLO 네트워크 아키텍처를 구성하였다.

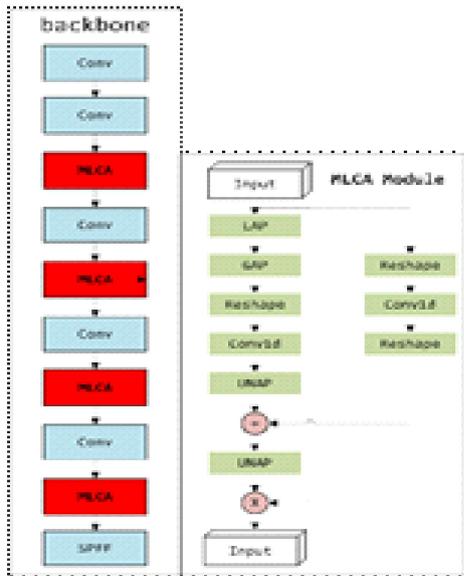


그림 1. MLCA를 적용한 YOLOv8n 모델의 구성도

3.3 실험 결과

모델은 5000장의 해상도 640x640의 새 이미지로 구성된 오픈 소스 데이터셋을 사용하여 학습되었다. 학습 과정 동안, 이미지 입력 크기를 640x640으로 설정하고, 초기 학습률을 0.01로 설정하였으며, 배치 크기를 64로 설정하였고, 총 300 에폭(epoch) 동안 학습을 진행하였다. 다음은 얻어진 결과이다.

평가결과 표 1에 나타난 바와 같이, 최신 YOLOv8n 모델에 비해 우리의 모델은 1%~1.5%의 성능 향상을 달성하

표 2. 향상된 YOLO모델의 성능 평가

Model	Layers	Parameters	FLOP	mAP
yolov5n	262	2.5M	7.2G	38%
yolov8n	225	3.1M	8.9G	40%
yolov10n	385	2.7M	8.4G	39%
Ours	255	3M	8.2G	41%

면서, 파라미터 수를 1K 줄이고 추론 복잡도를 0.7G 감소시켰다. 또한 그림 2에서 볼 수 있듯이 작은 크기의 새 또

한 정확하게 탐지하는 것을 확인 할 수 있으며 새와 유사한 비행기에서도 정확하게 새를 탐지하는 것을 확인 할 수 있다.



그림 2. 향상된 YOLO 모델을 이용한 조류 객체 탐지 결과

4. Conclusion

본 연구에서는 풍력발전소 운용시 조류의 접근을 차단하는 것을 목적으로하며, 새의 접근 유무에 조류 객체 탐지 모델을 사용한다. 기존 YOLO 모델에 MLCA기법을 적용하여 기존 YOLO 보다 높은 성능을 나타내는 모델을 제안하였다. 실험결과 특징 추출 아키텍처의 수정으로 모델의 정확도가 1%~1.5% 향상되었지만, 여전히 개선의 여지가 크다. 연구의 다음 단계에서는 더욱 정교한 피쳐 융합 기술과 새로운 탐지 헤드 기법을 조사하여 모델의 정확도와 추론 속도를 추가로 향상시키는 데 중점을 둘 것이다.

감사의 글

본과제(결과물)는교육부와한국연구재단의재원으로지원을받아수행된3단계산학연협력선도대학육성사업(LINK 3.0)의연구결과입니다. 또한, 본 연구성과는 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. RS-2024-00453747)

참고문헌

[1] Redmon, J., et al., You Only Look Once: Unified, Real-Time Object Detection. arXiv [cs.CV], 2016.
 [2] Dahang, W., et al., Mixed local channel attention for object detection. Engineering Applications of Artificial Intelligence, 2023. 123: p. 106442.

건설 XR 협업 서비스를 위한 웹 포털시스템

서명배, 박형진
한국건설기술연구원 미래스마트건설연구본부
e-mail : smb@kict.re.kr

Web portal system for construction XR collaboration services

MyoungBae Seo, HyungJin Park, SeongJong Oh
Korea Institute of Civil Engineering and Building Construction
Construction Automation Research Center

요 약

건설산업에 XR 기술을 기반으로 원격 협업 및 데이터 공유를 하기 위해서는 반드시 BIM(Building Information Modeling)을 기반으로 데이터를 다양한 장비에서 공유하는 서버 기반의 데이터 저장과 이를 관리하기 위한 기술이 매우 중요하다. 본 논문에서는 건설XR 협업 서비스를 위해 필요한 웹포털 서버의 핵심 기능을 정의하고 이를 기반으로 서버를 구성하고 웹 포털 트로토타입을 개발하였다. 웹 포털에서는 BIM 표준인 IFC 파일 관리방안, 향후 원격 협업을 위해 필요한 다양한 동기화 기능, 영상, 음성 공유 기능 등을 모두 고려하여 구성되었다. 향후 BIM 뷰어 등과 연동하여 다양한 이종장비에서 데이터를 공유하는 시스템에 확장 적용될 예정이다.

1. 서 론

건설산업에 XR(eXtended Reality) 기술을 적용하기 위해서는 다양한 중요한 요소가 존재하나 무엇보다도 건설 분야 3D 모델의 근간이 되는 BIM(Building Information Modeling) 데이터를 다양한 장비에서 손쉽게 조회하고 조회된 데이터를 공유하는 서버 기반의 데이터 저장 기술이 매우 중요하다. 더불어 이러한 BIM 데이터를 기반으로 이종기기간 협업을 위해서는 다양한 서비스 데이터를 저장하고 서비스할 수 있는 서버를 구축하고 본사에서 원격 XR 장비와 협업할 수 있는 본사용 BIM 모델 뷰어, 현장에서 BIM 모델 데이터를 정합하고 다양한 데이터를 입력하는 등 현장에서 활용가능한 XR 기반 App, 그리고 본사와 현장 데이터를 동기화 할 수 있는 기술, 다양한 협업이 가능한 기술 등의 구현이 필요하다. 본 논문에서는 서버를 기반으로 본사와 현장간 협업을 위해 건설XR 웹 포털시스템 구축방안 및 개발사례를 제시하고자 한다.

2. 이론적 고찰

2.1 XR과 건설 산업

XR(확장 현실)은 가상 현실(VR), 증강 현실(AR), 혼합 현실(MR)을 포함하는 포괄적인 기술로, 가상 환경과 실제 환경을 결합하여 사용자에게 몰입감 있는 경험을 제공한다. 건설산업에서는 XR을 기반으로 건축 설계 및 시공 계획 단계에서 시각적 피드백을 제공하여 설계 오류를 사전에 발견하고 수정하는데 활용하거나, AR을 사용하여 실제

현장에서 설계된 건축 요소를 실시간으로 확인할 수 있다. 또한 XR을 활용한 시뮬레이션을 통해 건설 작업자들은 안전 교육과 실제 작업 환경에 대한 훈련을 보다 효과적으로 받을 수 있다. 또한 MR 기술은 현장 관리자들이 가상의 모델과 실제 현장을 통합하여 시공 상태를 모니터링하고 문제를 해결하는 데 도움을 줄 수 있다.

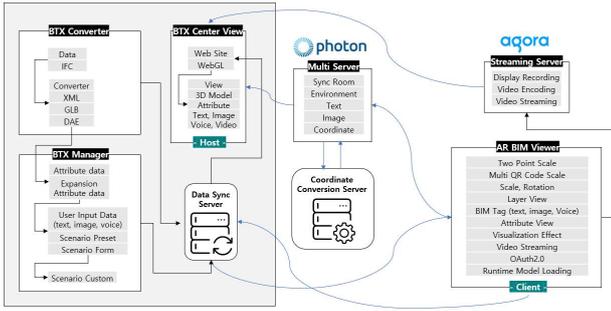
2.2 건설XR을 위한 웹 포털의 역할 및 기능 정의

건설 프로젝트는 다양한 전문가들이 참여하는 복잡한 작업으로 XR 기술을 활용하면 설계자, 엔지니어, 시공업체 등 다양한 이해관계자들이 실시간으로 정보를 공유하고 협력할 수 있다. 협력을 위해서는 데이터의 공유가 필수적이며 웹 포털 시스템은 모든 관련 정보를 중앙에서 관리하고 접근할 수 있는 플랫폼을 제공해야 한다. 이는 XR 기술로 생성된 가상 모델, 문서, 일정, 피드백 등을 포함한다. 웹 포털 시스템은 실시간 채팅, 화상 회의, 파일 공유 등의 협업 도구를 제공하여 팀원 간의 효과적인 소통을 지원할 뿐만 아니라 웹 포털에서 직접 가상 모델을 보고 수정할 수 있다. 또한 AR, VR, MR 콘텐츠의 통합 관리가 가능하고 다양한 디바이스에서 접근 가능하며, 사용자 권한 관리 및 데이터 보안을 통해 정보의 안전성을 보장해야 한다.

3. 건설 XR 협업 서비스를 위한 웹 포털시스템

3.1 건설XR 협업을 위한 서버 구축

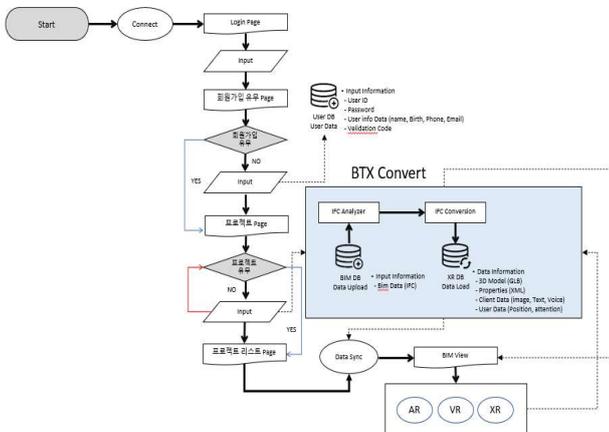
건설XR 기술을 기반으로 하는 다양한 건설 이해당사자 간 협업을 위해 PC, Mobile, 태블릿, AR 글래스 등을 활용한 데이터 공유가 필수적이므로 온라인상에서 데이터를 공유하기 위한 서버 구성이 필수적이다. 서버 구성을 위해 Web서버, DB서버, WAS서버, 통신서버, 데이터동기화 서버 등으로 구분하였다. Web서버는 Node.js, WAS 서버는 Tomcat을 활용하였으며, DB서버는 MySQL을 활용하였다. 협업을 위한 사용자의 위치정보 공유를 위해 게임엔진에서 많이 활용되는 Photon을 사용하였으며 실시간 영상 동기화를 위해서는 Agora를 활용하였다.



(그림 1) 서버 구성 및 서버간 데이터 연계도

3.2 웹서비스 포털 구축

웹 포털에서는 건설사별, 프로젝트별로 권한설정이 가능하고 프로젝트별로 BIM(본 논문에서는 BIM 표준인 IFC를 활용)을 포함한 다양한 도면을 저장한다. IFC 파일을 업로드하면 GLB, XML 등 형상, 속성으로 자동으로 분리되어 변환되고 데이터를 압축하는 모듈 역시 포털에 탑재되어 있어서 사용자가 XR 서비스 데이터로 변환할 필요가 없다. 또한 업로드된 IFC 파일을 분석하는 기능 등을 제공하며 PC에서 관리자가 다양한 도면을 향후 개발된 BIM 뷰어로 볼 수 있는 링크를 제공한다. 현장에서 XR 단말기로 동일한 현장의 3D 도면을 다운로드 하고 현장에서 발생하는 다양한 추가 데이터를 입력하게 되면 이 정보 역시 웹서비스 포털에서 데이터를 확인할 수 있게 구성하였다. 웹서비스 포털 시스템 개념도는 그림 2와 같다.



(그림 2) 웹 서비스 포털 개념도

이러한 개념을 근간으로 웹서비스 포털 프로토타입을 구축하였다. 웹에서 특정 아이피나 도메인을 치고 접근할 수 있다. 화면 구성을 살펴보면 왼쪽에는 사용자 최근 로그인 정보를 확인할 수 있고 건설사의 특성상 협력업체, 하청업체 등을 추가하며 사용자별 데이터 접근권한 등을 추가할 수 있다. 왼쪽 중단에 있는 “ifc프로젝트” 메뉴를 선택하면 프로젝트 명을 생성할 수 있고 IFC 파일 업로드를 하면 BIM to XR 컨버터가 동작되면서 XR로 서비스가 가능한 형태로 데이터가 변환된다. 프로젝트별 중단에 까만색 버튼은 향후 개발될 BIM 뷰어와 연동되게 구성하여 확장성을 고려하였다.



(그림 3) 웹 포털화면 및 IFC 파일 업로드 절차

4. 결론

본 논문에서는 BIM 데이터를 기반으로 건설협업 서비스가 가능한 웹 포털시스템을 설계하고 프로토타입 구현 사례를 제시하였다. 향후 다양한 XR 이종기기 연동 및 원격 협업이 가능한 기능을 추가할 예정이다.

감사의 글

이 논문은 과학기술정보통신부의 오픈XR(OXR) 융합연구사업(No. CRC21013)의 지원을 받아 수행된 연구임.

참고 문헌

- [1] 성균관대학교, 2019, “가상/증강현실 기반 공동주택 건설관리 선진화 기술개발 기획, pp.31~50
- [2] 조병욱, 조훈희, 손창백 (2021), 건설산업의 가상증강현실 활용실태 분석 및 개선방안, 대한건축학회 논문집 37(1), pp.229~237
- [3] 한국건설기술연구원, 2020, ”VR/AR 기반의 스마트 건설 가상화 시뮬레이션 기술 개발“, pp.46~49.

머신러닝을 활용한 홉의 알파산 함량 예측 모델 개발에 관한 연구

이현조¹, 백하영², 주원균^{3*}, 구현정^{4*}, 채철주^{5*}

¹한국농수산대학교 산학협력단, ²강원대학교 바이오기능성소재학과,

³한국과학기술정보연구원 데이터기반문제해결연구단,

^{4*}한국농수산대학교 작물·산림학부, ^{5*}한국농수산대학교 교양학부

e-mail : o2near@gmail.com, hjungkoo@korea.kr, bhy3014@naver.com, joo@kisti.re.kr, chae.cheoljoo@gmail.com

Study on the Development of a Hop Alpha Acid Content Prediction Model Using Machine Learning

HyunJo Lee¹, Hyun Jung Koo², Ha Young Baek³, Won-Kyun Joo^{4*}, Cheol-Joo Chae^{5*}

¹Industry-Academic Cooperation Foundation, Korea National University of Agriculture and Fisheries

²Dept. of Herbal Medicine Resource, Kangwon National University

^{3*}Dept. of Datacentric Problem Solving Research, Korea Institute of Science and Technology Information

^{4*}Dept. of Crops and Forestry, Korea National University of Agriculture and Fisheries

^{5*}Dept. of General Education, Korea National University of Agriculture and Fisheries

요약

본 논문에서는 기후변화가 홉의 생산량과 품질, 특히 알파산 함량에 미치는 영향을 분석하고, 머신러닝을 활용해 알파산을 예측하는 방법을 제안한다. 기후변화로 인해 주요 홉 생산국인 미국, 독일, 체코 등에서 생산량이 감소하고 있으며, 알파산 함량 또한 최대 32%까지 줄어들 것으로 예상된다. 국내 홉 생산량은 소비량의 10% 미만으로 대부분 수입에 의존하고 있으며, 가격 변동성이 크다. 기존 연구는 주로 홉의 연간 생산량에 집중했으며, 알파산 함량에 대한 장기적 연구는 부족하다. 그러므로 본 논문에서는 머신러닝을 이용하여 알파산 함량을 예측할 수 있는 방법을 제안한다. 홉의 알파산 함량 예측을 위해 kNN, DT, RF, XGB, LGBM, ADA 모델을 활용하였고, 각 모델의 장단점을 보완하며 예측 성능을 개선하였다. 실험 결과, 5가지 홉 품종에 대해 평균 오차율 0.3%의 높은 예측 정확도를 달성하였다.

1. 서론

기후변화는 농업기후 자원의 지리적·계절적 변화를 초래하여 작물 적응 지역의 변화, 토양 비옥도의 변화 등 농업 생산에 간접적인 영향을 주고 대기 중 CO₂ 농도 증가와 온도 상승은 작물의 광합성에 의한 물질 생산, 동화물질의 각 기관으로의 배분, 발육속도 등 작물의 생산성에 직접적인 영향을 미치고 있다. 이러한 기후변화에 따른 영향은 지리적 및 국지적으로 매우 다르게 나타날 것으로 예상되고 있다. 최근 연구에 따르면 기후변화로 인한 고온과 가뭄이 홉의 품질과 생산량에 영향을 미치고 있으며, 주요 홉 생산국에서는 생산량이 감소하고 있다. 주요 홉 생산국의 생산자 협회에 따르면 2022년 미국 내 홉 생산량은 전년 대비 12% 감소하였으며 같은 기간 독일은 21%, 체코는 40% 이상 홉 생산량이 감소하였다. 또한 맥주의 품질을 결정하는 알파산 함량은 13~32%까지 감소할 것이라고 예상하고 있다. 또한 홉은 알파산, 베타산, 플라보노이드 등 다양한 유효성분에 대한 활용 방안을 연구하고 있으나 국내 홉 생산량은 전체 소비량의 10% 미만으로, 대부분 해외 수입에 의존하고 있다. 기후변화 등으로 인해 해외 홉 생산량 변동, 환율 변동 등으로 인한 홉 가격 변동성이 높기 때문에 홉의 기능성 및 활용 분야에 대한 연구 개발이 필요한 실정이다. 그러므로 본 논문에서는 머신러닝을 이용하여 홉의 알파산을 예측

할 수 있는 방법을 제안한다.

2. 관련 연구

홉에 대한 대부분의 연구는 연간 생산량에 대한 데이터에 중점을 두었으며, 알파산 함량 및 페놀 함량과 같은 지표를 다루는 연구는 많이 이루어지고 있지 않다. 미국의 경우 농무부가 제공한 National Hop Report에서는 홉의 단위 면적당 생산량, 판매 가격, 재배 면적, 지역별 생산량 등의 정보를 제공하지만, 홉 품질에 관한 데이터는 포함하고 있지 않는다. 유럽의 경우는 홉 생산 및 품질 특성에 대한 연구가 활발히 이루어지고 있으며, 특히 체코를 중심으로 매년 홉의 생산량 및 품질 특성에 대한 보고서가 발간되고 있으며, 2021년부터 2023년까지 체코산 홉의 알파산 함량을 분석하였다[1]. 그러나 해당 연구는 데이터 수집 기간이 짧아 장기적인 경향성을 파악하는 데에는 한계가 있다. Martin Mozny 등은 1970년부터 2018년까지 독일 3개 지역, 체코 1개 지역, 슬로베니아 1개 지역의 홉 생산량 및 알파산 함량 데이터를 분석하였으며, 해당 데이터를 공개하여 장기적인 변화 추이를 평가할 수 있는 중요한 자료를 제공하고 있다 [2]. 본 논문에서는 머신러닝을 이용하여 홉의 알파산을 예측할 수 있는 방법을 제안한다.

3. 머신러닝을 이용한 홉의 알파산 예측 방법 및 결과

본 논문에서는 홉의 알파산 함량 예측 성능을 향상시키기 위해 앙상블 기법을 사용하였다. 앙상블 기법은 최근 단일 학습자보다 일반화 및 견고성 측면에서 우수한 성능을 보이며 다양한 분류 및 회귀 문제에 사용되고 있다. 앙상블 기법은 서로 다른 다양한 모델을 층을 쌓듯이 조합하여 새로운 모델을 만드는 방식으로 베이스와 메타의 두 단계를 거치면서 각 모델들의 장점은 취하고 약점은 보완하면서 최종 모델의 성능을 향상시킨다. 알파산 예측을 위해 1970년 ~ 2018년 데이터를 사용하였으며, [그림 1]은 실험에 사용한 데이터의 일부를 보여주고 있다.

Yields	Zatec	Spalt	Hallertau	Tettnang	Celje	Alpha	Zatec	Spalt	Hallertau	Tettnang	Celje
1970	1.19	1.75	2.135	1.995	1.49	1970	5.4	5.3	5.5	4.8	6.7
1971	0.805	1.45	1.565	1.765	1.055	1971	4.6	4.5	4.5	4.7	5.8
1972	1.19	1.42	1.66	1.86	1.18	1972	5.2	4.8	5.3	4.2	6.2
1973	1.12	1.75	1.95	1.4	1.28	1973	5.6	5.5	5.7	4.8	5.6
1974	0.76	1.44	1.67	1.48	1.2	1974	4.75	4.4	4.6	3.45	5.95
1975	1.06	1.51	1.61	1.33	1.18	1975	5.35	5.4	5.25	5.1	5.7
1976	0.92	1.4	1.45	1.18	1.12	1976	3.9	3.8	4	2.8	4.6
1977	1.14	1.55	1.98	1.37	1.24	1977	4.9	5.1	4.8	3.9	5.4
1978	0.92	1.79	1.73	1.44	1.37	1978	3.8	3.9	4	3.4	6.3
1979	1.08	1.77	1.83	1.48	1.29	1979	4.3	4.8	4.5	3.9	6
1980	1.14	1.26	1.55	1.05	1.48	1980	3.9	3.8	3.9	4	3.8
1981	1.14	1.68	1.39	1.56	1.73	1981	4.3	4.7	4.8	4.2	5.3
1982	1.06	1.73	1.58	1.55	1.72	1982	3.7	4.2	3.7	3.9	5.6
1983	1.05	1.36	1.36	1.36	1.35	1983	2.9	3	3.1	2.9	4.2
1984	0.91	1.37	1.19	1.38	1.79	1984	4.2	4.2	4.1	3.9	6.5
1985	0.96	1.22	1.22	1.39	1.62	1985	3.5	3.9	3.7	3.4	5.8
1986	0.82	1.32	1.08	1.32	1.43	1986	4.3	4.8	4.8	3.7	5.4
1987	0.9	1.22	1.09	1.27	1.75	1987	5.3	5.2	5.7	3.6	5
1988	1.34	1.12	1.02	1.27	1.5	1988	4.1	4.2	4.3	3.6	4.8
1989	0.97	1	1.09	1.26	1.29	1989	4.2	4	4.2	3.3	5.7
1990	0.84	0.91	0.83	1.14	1.45	1990	4.3	3	3.6	3.7	5.6
1991	0.77	1.2	1.35	1.34	1.55	1991	5	4.5	4.6	3.9	5.5
1992	0.99	0.99	0.86	1.1	1.54	1992	3.9	3.5	3.6	3	3.2
1993	0.9	0.99	1.28	1.28	1.04	1993	4.2	4.1	4	3.4	3.6
1994	0.9	1.18	0.96	1.19	1.28	1994	2.6	2.8	2.9	2.4	2.8
1995	0.94	1.05	0.7	1.03	1.39	1995	3.3	3.3	2.6	2.6	2.9
1996	1.02	1.31	0.79	1.29	1.28	1996	5.3	5.4	4.6	3.8	4.3
1997	0.92	1.01	1.17	1.4	1.44	1997	5.1	5.2	5	3.8	4.4
1998	0.84	0.83	1.17	1.25	1.15	1998	4.6	4.1	3.8	3.6	4.3

(그림 1) 홉 생산량 및 알파산 데이터

홉의 알파산 함량 예측을 위해 사용된 베이스 모델로는 kNN, DT, RF, XGB, LGBM, ADA를 사용하였으며, 각 모델별 하이퍼파라미터 값은 <표 1>과 같이 설정하였다. 제안 방법을 이용하여 5가지 품종에 대한 알파산 함량을 예측한 결과 <표 2>와 같이 평균 오차율이 0.3%로 높은 예측 결과를 도출하였다.

(표 1) 제안 모델에 사용된 파라미터값

Model	Hyperparameters		Optimized Hyperparameter Values
	Name	Search Range	
kNN	n_neighbors	3 to 10, step = 1	3
	P	1.0 to 2.0, step = 1.0	1.0
	leaf_size	1 to 100, step = 1	2
DT	criterion	{'squared_error', 'poisson', 'friedman_mse', 'absolute_error'}	squared_error
	max_depth	1 to 10, step = 1	10
	max_leaf_nodes	2 to 10, step = 1	10
	min_samples_split	2 to 10, step = 2	8
	min_samples_leaf	1 to 4, step = 1	1
RF	n_estimators	10 to 2000, step = 1	361
	max_features	{'auto', 'sqrt'}	auto

	max_depth	1 to 16, step = 1	10
	min_samples_split	2 to 10, step = 2	10
	min_samples_leaf	1 to 4, step = 1	4
XGB	n_estimators	10 to 2000, step = 1	1024
	max_depth	1 to 20, step = 1	20
	learning_rate	0.01 to 1.00, step = 0.01	0.3
LGBM	alpha	0.01 to 1.00, step = 0.01	0.01
	n_estimators	10 to 2000, step = 1	1877
	max_depth	1 to 20, step = 1	10
	min_child_weight	1 to 300, step = 1	242
	min_child_samples	10 to 50, step = 1	49
	learning_rate	0.01 to 1.00, step = 0.01	0.41
	n_estimators	10 to 2000, step = 1	1582
ADA	learning_rate	0.01 to 1.00, step = 0.01	0.82

(표 2) 홉 알파산 예측 실험 결과

	Celje	Hallertau	Spalt	Tettnang	Zatec
LinearR	0.645	0.620	0.574	0.405	0.505
kNN	0.677	0.605	0.549	0.402	0.524
SCM	0.614	0.561	0.527	0.378	0.470
DT	0.397	0.417	0.362	0.236	0.349
RF	0.359	0.416	0.369	0.249	0.345
XGB	0.003	0.001	0.003	0.001	0.001
ADA	0.413	0.474	0.439	0.301	0.387
LGBM	0.176	0.066	0.055	0.149	0.078
CAT	0.014	0.007	0.003	0.001	0.002

4. 결론 및 향후 연구

본 논문에서는 1970년부터 2018년까지 독일 3개 지역, 체코 1개 지역, 슬로베니아 1개 지역의 홉 생산량 및 알파산 함량 데이터를 머신러닝을 이용하여 알파산 함량을 예측할 수 있는 방법을 제안하였다. 제안 방법을 이용한 결과 평균 오차율 0.3%로 높은 예측 성능 측정 결과를 보였으며, 제안 방법을 통해 기후변화에 따른 홉의 효율적인 재배에 기여할 수 있다. 향후 연구로는 홉의 알파산과 베타산의 함유량 및 비율을 향상에 기여할 수 있는 요인을 분석하고 예측할 수 있는 인공지능기반의 시스템을 개발할 예정이다.

ACKNOWLEDGMENTS

이 논문은 2024년도 한국과학기술정보연구원(KIST)의 지원을 받아 수행된 연구입니다. (과제번호: (KISTI)J24JR008-24)

참고 문헌

- [1] KROFTA, Karel, et al. Alpha acids content in Czech hops from the harvest of 2021-forecast, reality, trends. Kvasny prumysl, 2022, 68.1: 564-571.
- [2] MOZNY, Martin, et al. Climate-induced decline in the quality and quantity of European hops calls for immediate adaptation measures. Nature Communications, 2023, 14.1: 6028.

머신 러닝 기반 넙치 질병 분류 알고리즘

정중호*, 김찬진*, 손현승**, 임한규***, 최한석*
 빛가람정보(주)*, 국립목포대학교 컴퓨터공학과**,
 국립목포대학교 해양수산자원학과***

wkrlsk23@bgrinfo.co.kr, aqs450@bgrinfo.co.kr, hson@mnu.ac.kr,
 limhk@mokpo.ac.kr, chs@bgrinfo.co.kr

Flatfish disease classification algorithm based machine learning

Jeong JongHo*, Kim ChanJin*, Son HyeonSeung**, Lim HanKyu***,
 Choi HanSuk*

Bitgaram Info Co., LTD.*,

Department of Computer Engineering, Mokpo National University**,
 Department of Marine and Fisheries Resources, Mokpo National University***

요약

본 논문에서는 넙치 양식 산업의 생산성 향상을 위해 머신러닝 기반 넙치 질병 분류 알고리즘을 제안하고 그 성능을 분석하였다. 넙치 질병 데이터셋을 구축하고, SGDclassifier 모델과 GridSearchCV를 활용하여 질병 분류 모델을 학습하였다. 2023년과 2024년 넙치 데이터셋을 사용하여 모델 성능을 평가한 결과, 두 데이터셋 모두 Top-3 정확도 80% 이상을 달성하며 효과적인 질병 분류 성능을 보였다. 특히, 2024년 데이터셋은 더 많은 증상 정보를 포함하여 더 높은 정확도를 나타냈다. 이는 증상 정보가 많을수록 모델의 질병 분류 성능이 향상됨을 시사한다. 향후 연구에서는 더 많은 데이터셋과 가치를 적용시키는 전처리 기법등을 통해 모델 성능을 더욱 향상시킬 계획이다.

1. 서론

양식 넙치의 질병 관리는 생산성 향상에 있어 중요한 요소이며, 질병 증상을 조기 발견하고 정확히 분류하는 것은 넙치 생산성 향상을 위한 중요한 과제이다. 이전 연구에서는 Vitdet, YOLO 등 다양한 모델을 활용하여 넙치 이미지에서 질병 증상을 감지하는 성능을 비교하였다.[1] 본 논문에서는 감지된 증상을 기반으로 넙치 질병을 효과적으로 분류하는 알고리즘을 제안한다.

알고리즘 실험 결과 2023년 데이터셋과 2024년 데이터 셋 모두 Top-3 내 80% 이상의 좋은 성능을 보여주었으며, 이는 넙치 데이터셋을 질병 분류 알고리즘이 효과적으로 분류하였음을 보여주었다.

2. 넙치 데이터

2.1 넙치 데이터 클래스 코드

넙치 질병 분류를 위한 데이터셋 구축을 위해 질병 병원체별로 나타나는 부위별 증상을 코드화하여 라벨링을 실시한 뒤 데이터셋에 저장하였다. 본 연구에서는 넙치 양식장에서 주로 발생하는 에드워드병, 비브리오팀 등 총 7 가지 질병을 각각 고유한 영문자 2단어 코드로 정의하였다. 또한 질병이 발생하는 17 부위는 2단어 코드로, 23종류의 증상은 1단어 코드로 정의하였다. 이를 통해 각 질병 병원체에 대한 부위와 증상을 조합한 5자리 클래스 코드를 생성하여 데이터 라벨링에 사용하였다.

2.2 넙치 데이터셋

넙치 질병 분류를 위한 학습 데이터는 넙치 이미지와 해당 이미지의 질병 및 증상 정보를 담은 라벨링 데이터로 구성된다. 본 논문에서 사용되는 넙치 데이터는 아래 그림과 같은 구성으로 제작되었다. 상단 테이블은 각 개체의 인덱스, 증상, 질병을 나타내며 하단 테이블은 증상의 유무를 이진 벡터로 나타낸 증상 테이블이다. 각 행은 개체에 해당하고, 각 열은 특정 증상의 존재 여부를 나타낸다. 이러한 구조를 통해 각 개체의 증상들을 하나의 세트로 묶어 머신 러닝 모델이 질병을 예측하거나 분류하는데 활용될 수 있도록 하였다.

Index	Symptoms	Disease
JCS231201	...	EL
JAI230948	...	MA
⋮	⋮	⋮
JCS230501	...	PO

	ASC	ASH	ASM	...	REM	SPM
○	x	x	x	...	x	○
⋮	⋮	⋮	⋮	⋮	⋮	⋮
x	x	○	...	○	○	

Symptoms table

(그림 1) 넙치 질병 데이터셋 구성도

3. 질병 분류 알고리즘

3.1 머신러닝 기반 넙치 질병 분류 알고리즘

데이터를 특정 기준에 따라 여러 그룹으로 나누는 작업을 분류라고 하며, 이때 분류 알고리즘을 사용한다. 본 논문에서는 넙치 질병 분류를 위해 머신러닝 기반 분류 알고리즘을 사용한다

넙치 질병 데이터는 질병 정보가 라벨링된 데이터를 활용하여 데이터셋을 구축하고, 이를 지도 학습 방식으로 학습시켜 질병을 분류한다. 지도 학습은 입력 데이터와 해당 데이터의 정답 라벨이 쌍으로 구성된 데이터셋을 사용하여 모델을 학습시키는 방법이다. 이를 통해 입력 데이터의 특징과 라벨 간의 관계를 학습하고, 새로운 데이터가 주어졌을 때 해당 데이터의 라벨을 분류할 수 있다.

질병 분류를 위해 SGDclassifier 모델을 사용하여 데이터를 학습하였다. SGDclassifier는 대용량 데이터를 빠르게 학습하며, 무작위로 선택된 데이터 일부를 사용해 모델을 점진적으로 개선하는 머신러닝 기반 분류 알고리즘으로, 넙치 질병 분류에 적절하다.[2] 또한, GridSearchCV를 사용하여 모델의 하이퍼파라미터 최적화를 수행하였다.[3]

4. 실험 및 분석

질병 분류 알고리즘의 성능을 검증하기 위해 2024년 라벨링된 데이터를 사용하여 학습 데이터셋을 구축하였다. 모델의 일반화 성능을 평가하고 과적합을 방지하기 위해 데이터셋을 훈련용 80%와 테스트용 20%로 분할하였고 k-fold 교차 검증 기법을 적용하였다.[4] 테스트 데이터셋은 2023년 라벨링된 데이터셋과 2024년 라벨링된 데이터셋을 각각 사용하였다.

(표 1) 테스트 데이터셋 질병별 구성

항목	2023년 테스트 데이터셋	2024년 테스트 데이터셋
평균 증상 갯수	3개	7개
여윍증	2개체	5개체
에드워드병	5개체	5개체
스쿠티카병	6개체	5개체
연쇄구균병	4개체	5개체
활주세균병	5개체	5개체
패혈증	5개체	5개체
비브리오증	5개체	5개체
정상	5개체	5개체

(표 2) 테스트 데이터셋 학습 결과

model	Top-1	Top-2	Top-3
2023년 데이터	41%	68%	81%
2024년 데이터	75%	85%	95%

2023년 데이터와 2024년 데이터의 머신러닝 분류 알고리즘의 성능을 평가하기 위해 Top-N 방식을 사용하였다. Top-N 방식은 예측 결과 중 상위 N개 안에 정답이 포함되었는지를 확인하여 모델의 성능을 측정하는 방법이다.

학습 결과, 2023년 데이터셋은 Top-1부터 Top-3까지 정확도가 각각 41%, 68%, 81%로 나타났고, 2024년 데이터셋은 Top-1부터 Top-3까지 정확도가 각각 75%, 85%, 95%로 나타났다. 두 데이터셋 모두 Top-3까지 80% 이상의 정확도를 보이며 머신러닝을 통한 질병 분류에서 좋은 성능을 보여주었다.

특히, 2024년 데이터셋의 결과가 2023년 데이터셋보다 우수하였는데, 이는 2024년 데이터셋이 평균적으로 더 많은 증상을 포함하고 있어 모델이 질병 패턴을 쉽게 파악할 수 있었기 때문으로 판단된다. 넙치 데이터셋의 특징을 고려하면, 증상 정보가 많을수록 모델의 질병 분류 성능이 향상 되는 것을 확인할 수 있었다.

5. 결 론

본 논문에서는 넙치 양식 산업의 생산성 향상을 위해 머신러닝을 사용한 분류 알고리즘을 소개한다. 넙치 라벨링 데이터를 수집 년도를 나누어 케이스를 제작하였고 각각의 결과를 확인하여 성능을 비교하였다.

실험 결과, 2023년 데이터셋은 Top-1 정확도 41%, Top-2 정확도 68%, Top-3 정확도 81%를 기록하였고 2024년 데이터셋은 Top-1 정확도 75%, Top-2 정확도 85%, Top-3 정확도 95%를 기록하였다. 2024년 데이터셋이 2023년 데이터셋에 비해 모든 Top-N 정확도에서 높은 성능을 보였으며 두 데이터셋 모두 Top-3 정확도가 80%를 넘어 머신러닝 모델이 질병 분류를 효과적으로 하였음을 볼 수 있었다. 향후 연구에서는 더욱 많은 데이터셋을 학습 시키고 가중치를 적용시키는 등의 전처리를 통해서 성능 향상을 확인해볼 예정이다.

감사의 글

이 논문은 2022년도 해양수산부 재원으로 해양수산과학기술진흥원의 지원을 받아 수행된 연구임(20220579, 빅데이터 기반 양식생산성 향상기술).

참고 문헌

- [1] 조경원 외 6명. “넙치질병증상분류를위한 객체탐지 딥러닝 모델 성능 평가”, 『한국스마트미디어저널』 Vol.12 No.10, 2023.11.
- [2] SGDclassifier documents, https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
- [3] GridSearchCv documents, https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html
- [4] K-fold 교차검증, <https://bbdata.tistory.com/10>

인공지능의 활용이 세금신고업무에 미치는 영향

장성일, 김정연

단든세무회계, 상명대학교

e-mail : dandeunj@gmail.com, jykim@smu.ac.kr

The Impact of Artificial Intelligence on Tax Reporting

Seongil Jang, JeongYeon Kim

Dandeun tax accounting, Dept. of Business Administration, Sangmyung University

요약

알파고 이후 ChatGPT에 이르기까지 현재 AI를 활용하여 각 분야별 효율성을 높이기 위한 노력이 진행 중이다. 이러한 인공지능을 활용할 때 일반소비자들에게 미치는 영향을 확인하고자 한다. 세법에 관한 전문적인 지식이 없는 일반 납세자를 대상으로 인공지능 서비스를 활용하여 사업소득금액을 계산하는 것이 가능한지 시도해 보았다. 또한 세법에 관한 전문적인 지식이 있는 세무사가 주어진 수입 금액과 지출내역을 토대로 ChatGPT를 활용하여 사업소득금액을 계산해 보았다. 아직 우리 사회에 인공지능을 활용한 업무처리가 보편화된 것이 아닌 상황에서, 학습을 통해 고도화된 모델을 지속적으로 발전시켜야 하겠지만 동시에 이를 활용하는 소비자 측면에서도 인공지능의 특성을 이해하고 이를 보다 효과적으로 활용할 수 있는 방안을 모색할 필요가 있어 보인다.

1. 서론

인공지능에 대한 사회적 관심이 높아지면서 기업 및 공서에서는 이를 이용하는 사업모델에 집중하고 있다. 반면 일반인들은 업무효율성이나 개인적인 관심정도에 따라 다양한 반응을 보이고 있다. 따라서 세금신고라는 공통의 분야에서 일반인이 인공지능을 얼마나 활용할 수 있는지 확인해 보고자 한다. 일반사용자의 접근성을 고려하여 인지도가 높고 무료로 제공되는 ChatGPT를 활용하였고, 세무적 복잡성을 고려하여 사업소득이 있는 간편장부대상자를 대상으로 하였다. 또한 비교가능성을 높이고자 최종세액계산이 아닌 사업소득금액을 계산하는 것을 목표로 하였다.

2. 본론

일반인이 ChatGPT를 활용하여 사업소득금액을 계산하였을 때 배경지식의 유무 및 정보를 취득하는 능력, 절세하고자 하는 의지, 정보에 대한 비판적 수용 등 개인적인 요소에 따라 활용성에 차이가 있었다.

평소 세금에 관심이 많았던 30대 여성(상품판매업 사업소득과 근로소득 발생)은 본인 사업의 수입금액의 파악 및 관련된 경비를 나름대로 정리 해두었을 정도로 기본적인 배경지식 있는 상태였다. 이를 토대로 ChatGPT를 통해 사업소득금액을 계산하는 것은 어렵지 않았다. 다만, 주어진 수입금액과 경비만 차감하였을 뿐 질문자가 절세에 대한 추가 질문을 하지 않았기 때문에 사업소득금액의

계산이 더 정교해지지는 않았다. 질문자의 경우 상품매입가, 운송비 등 주요경비에 대한 배경지식은 있었지만 기부금, 지급수수료 등에 다양한 비용구조에 대한 내용은 모르고 있었다. ChatGPT활용 할 때 절세를 더 시도하지 않은 이유에 대해서는 사업소득이 부업일 뿐이고 질문 당시 본인이 생각하는 비용만 잘 반영했다면 크게 신경쓰지 않았다고 답변했다.

근로소득만 있던 30대 남성의 경우 프리랜서 사업소득이 발생하였는데, 세금에 대한 지식이 전무했기 때문에 소득금액을 계산하는 방법 및 필요경비의 종류부터 ChatGPT에 질문하였다. 이를 토대로 본인의 필요경비를 집계하여 사업소득금액을 계산한 결과 세부담을 크게 느꼈고 추가 질문을 통해 세부담을 줄이고자 하였다. 그러나 질문자는 필요경비와 소득공제, 세액공제의 구분을 할 수 없어 필요경비에 대한 질문이 아닌 각종 공제에 대한 질문을 하였다. 따라서 사업소득에서는 필요경비에 해당하는 보험료, 기부금을 소득공제 및 세액공제로 반영하는 것으로 착각하고 있었다. 다만 질문자의 경우 ChatGPT를 활용하는 과정에서 답변결과에 대한 신뢰가 떨어져 실제 세금신고를 할 때 이 질문과 답변 내용을 토대로 세무사에게 재차 확인하는 과정을 거칠 것이라 답변했다.

사업소득만 있는 30대 남성(제조업)의 경우 세무사를 통한 종합소득세 신고경험이 있었고 지출의 사업관련성이 중요하다는 것을 인지하고 있었다. 따라서 지출내역의 필

참 고 문 헌

요경비 인정여부를 직접 묻는 등 질문방식의 차이가 있었고 보다 구체적인 답변을 얻어 사업소득금액을 더 정교하게 계산할 수 있었다. 다만, 복잡하고 민감한 문제에 대해서는 전형적인 답변의 반복이어서 만족스럽지 않았다고 했다.

세무사가 주어진 자료를 토대로 ChatGPT를 활용하여 사업소득금액을 계산해 본 결과 세무사가 사전에 직접 계산한 금액과 결과값은 일치했다. 비교적 간단한 수준의 사업소득금액 계산이었고, 세무사는 사업관련성에 입각해 다양한 비용구조에 대한 이해가 있었기 때문이다. 또한 질문을 할 때 이미 답을 알고 있기 때문에 ChatGPT가 부족하거나 잘못된 답을 했을 때 추가질문을 하여 ChatGPT의 답변을 보다 정교하게 보완할 수 있었다.

일반인이 ChatGPT를 활용하여 사업소득금액을 계산하고자 할 때, 분석대상이 되는 자료의 정제 정도, 답변의 적정성에 따른 세무리스크 등 여러 변수들과 윤리적, 법적인 문제가 있으나, 일반이용자의 경우 결과를 정교화하는 추가적인 질문을 효과적으로 하지 못하거나 개선의 여지가 있음에도 이를 인지하지 못하고 마무리 지어 절세측면에서 효과성이 다소 떨어졌다. 그러나 이것은 추가적인 역량을 소모하는 작업이므로 개인의 성향, 목적에 따라 그대로 세금신고를 마치거나 전문가 상담 전 참고자료 등으로 업무이용범위를 달리 삼을 수 있을 것이다.

참고로 국세청 홈택스에서 제공하는 AI세금비서의 경우 하단에 키워드를 제공한 후 예상질문을 제시하고 관련 질문에 대한 답변을 선제적으로 제공하고 있다. 이는 정교한 결과를 얻기 위한 질문에 소모되는 역량을 줄여주는 분명한 이점이 있지만 개별 특성을 모두 고려할 수는 없는 한계도 있다.

3. 결 론

결과에 대한 정교함을 위해서는 결국 이용자의 배경지식이 풍성해야 하는 것이 자명하지만, 일반인 입장에서는 개인의 의지와 더불어 답변에 대해 비판적인 수용을 하는 것이 중요한 요소로 보인다. 실제 필요경비에 대해 자세히 요구하는 질문을 동일하게 반복할수록 보다 자세하고 구체적인 답변을 얻을 수 있었고 더 절세할 수 있는 방법을 요구하는 질문을 하는 것으로 다른 고려사항을 발견하여 보다 정교한 결과에 이를 수 있었다.

Acknowledgement

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음” (IITP-2024-RS-2024-00438056)

- [1] 이창규. “인공지능(AI) 기술의 세무업무 활용과 과제”, 『한국조세법학회 조세논총』 제8권 제4호, 2023. pp.143-169.
- [2] 국세청 보도자료. 2023년 세금비서 서비스 도입 1년간의 성과, 2023.12.31.
- [3] 서울신문.세금 조회·납부도 AI로 한다… 국세청, 내년 ‘AI 홈택스’ 개통.24.05.21. <https://www.seoul.co.kr/news/economy/2024/05/21/20240521500141>
- [4] 설민수. “머신러닝 인공지능의 법 분야 적용의 현재와 미래:미국의 현황과 법조인력 구조 및 법학교육에 대한 논의의 중심으로”, 『한국법학원 저스티스』 통권 제156호, 2016.10. pp.269-302.
- [5] 조상규. “인공지능 세무대리 프로그램의 법적책임”, 『중앙법학회 중앙법학』 제19집 3호, 2017.09. pp.67-94.
- [6] 오수경,김제이,Marilou Hongliangko,강채현,오수연,정수정,임다미. “인간-AI 협업에 영향을 미치는 요인:주제범위 문헌고찰을 바탕으로”, 『멀티미디어학회 멀티미디어학회 논문지』 제27권 제8호, 2024.08. pp.945-9958.

CUDA Core와 Tensor Core를 사용한 두 개의 소규모 행렬의 곱 연산 성능 비교

김인서, 김진성

중앙대학교 컴퓨터공학과

e-mail : inseo764@cau.ac.kr, kimjsung@cau.ac.kr

Performance Comparison of Matrix Multiplication Using CUDA Core and Tensor Core for Small Matrices

Inseo Kim, Jinsung Kim

School of Computer Science and Engineering, Chung-Ang Univ.

요 약

LLM이 급격히 발전하면서 행렬 간 곱셈의 중요성은 점점 증대되고 있는데, 행렬 간 곱셈 연산은 상당히 계산집약적이라는 특징을 띤다. 따라서 다양한 방식으로 이 계산을 병렬화하고자 하는 연구가 진행되어왔다. 이때, GPU에서 일반적인 계산 유닛을 사용해서 병렬화를 했을 때도 행렬간 곱셈 연산에 필요한 시간은 크게 줄어들지만, NVIDIA GPU에서는 NVIDIA Volta architecture부터 이 연산을 더 가속화할 수 있는 연산 유닛인 Tensor core를 도입하여 더 높은 성능 향상을 이루고자 하였다. 하지만 Tensor core로는 한 번에 특정 크기의 행렬만 계산할 수 있다는 제약이 있다. 그래서 큰 행렬의 연산을 할 때는, 큰 행렬을 Tensor core에서 지원하는 특정 크기로 잘라서 Tensor core를 사용하여 연산하고, 그 값을 누적하는 방식을 사용한다. 이러한 상황에서 Tensor core에서 지원하는 크기보다 작은 행렬의 곱을 더 빠르게 계산하는 방식이 무엇인지 알아보려고 Cuda Core를 사용하는 방식과 패딩을 덧붙여 Tensor core를 사용하는 방식 두 가지로 구현하였고 그 성능을 비교하고자 하였다. 실험 결과 padding 과정을 병렬화하지 않은 경우는 일반 cuda core를 사용한 커널보다 느린 성능을 보였으며, padding 과정을 병렬화한 경우 두 배 이상 빠른 성능을 보였다. 또한 행렬의 곱셈을 연산하는 부분보다 CPU의 데이터를 GPU로 옮기는 시간이 가장 많은 비중을 차지하는 것으로 보아 해당 과정의 성능을 향상시키는 것도 성능에 결정적인 영향을 미칠 수 있음을 알 수 있다.

1. 서론

수 많은 LLM 서비스와 상향평준화된 컴퓨터 그래픽의 품질 등에 필요한 행렬 간 곱셈의 중요성은 점점 증대되고 있는데, 이 연산은 상당히 계산집약적이라는 특징을 띤다. $M \times K$ 사이즈 행렬 A, $K \times N$ 사이즈 행렬 B를 곱하는 연산을 해서 $M \times N$ 사이즈 행렬 C를 구하는 연산은 총 $2MKN$ 번의 연산이 필요하다. 이러한 연산을 순차적으로 처리한다면, 상당한 부담이 될 수 있다.

하지만, 만약 GPU를 사용하고, 한 thread당 행렬C의 요소 중 한 개를 계산하도록 시킨다면, 한 번의 연산에 1만 킬로바이트의 시간이 소요된다고 가정할 때, 2K만킬로바이트의 시간이 소요된다. 만약 이를 일반적인 직렬 컴퓨터에서 실행한다면 2MKN만킬로바이트의 시간이 필요했을 것이다.

이렇듯 GPU에서 일반적인 계산 유닛을 사용해서 병렬화를 했을 때도 행렬간 곱셈 연산에 필요한 시간은 크게 줄어들지만, NVIDIA GPU에서는 이 연산을 더 가속화할 수 있는 연산 유닛인 Tensor core를 도입하여 더 높은 성능 향상을 이루고자 하였다. 하지만 Tensor core로는 한 번에 특정 크기의 행렬만 계산할 수 있다는 제약이 있다. 그래서 큰 행렬의 연산을 할 때는, 큰 행렬을 Tensor core

에서 지원하는 특정 크기로 잘라서 Tensor core를 사용하여 연산하고, 그 값을 누적하는 방식을 사용한다.

그렇다면 Tensor core에서 지원하는 크기보다 작은 행렬의 곱은 어떻게 계산하는 것이 효율적인지에 대해 알아보려고 나는 이 연산을 CUDA core를 사용하는 방식과 패딩을 덧붙여 Tensor core를 사용하는 방식 두 가지로 구현하였고 그 성능을 비교하고자 하였다.

2. 관련 연구

2.1 Tensor core

Tensor core는 기존 CUDA core를 사용한 행렬 곱 연산보다 더 빠른 연산을 위해서 NVIDIA Volta GPU architecture부터 도입되었다. Tensor core는 행렬 곱 연산에 있어서 몇 가지 특정 자료형과 그에 따른 fragment 사이즈를 지원한다. 또한 Tensor core는 warp 단위로 그 연산이 이루어진다는 특징이 있다.

3. 소규모 행렬 곱셈

3.1 실험 환경

Windows 11의 Windows Subsystem for Linux(WSL)

환경에서 CUDA driver 12.4 버전을 사용하였으며, CUDA compute capability는 8.6을 사용하였다. 또한 하드웨어는 NVIDIA GeForce RTX 3090과 11th Gen Intel(R) Core(TM) i9-11900K @ 3.50GHz를 사용하였다.

3.2 실험 방법

텐서 코어에서 지원하는 행렬의 크기보다 작은 사이즈의 행렬 곱셈 연산(A x B = C)을 할 때 CUDA core 혹은 tensor core를 사용하는 경우 성능의 차이를 비교하기 위해 총 3가지 방식으로 행렬 곱셈을 구현하였으며, (m, n, k)값이 (1, 1, 1), (2, 2, 2), (4, 4, 4), (8, 8, 8), (16, 16, 16)인 경우 실행 시간을 측정하여 비교하였다. 또한 Nsight Systems를 이용하여 커널의 동작을 시각적으로 분석하였다.

첫 번째 커널 gemm_cudacore에서는 thread 1개당 행렬 C의 1개의 element를 계산하도록 하여 행렬 곱셈 과정을 병렬화 하였으며, 두 번째(gemm_tensorcore)와 세 번째(gemm_tensorcore_2) 커널의 경우는 NVIDIA CUDA API의 wmma API를 사용하여 구현하였으며, 자료형은 half를 사용하였고, Tensor core의 fragment의 사이즈는 (m, n, k) = (16, 16, 16)을 사용하였다.

두 번째 커널(gemm_tensorcore)과 세 번째 커널(gemm_tensorcore_2)의 차이는 padding을 넣어주는 방식에 있다. 두 번째 커널의 경우는 padding을 호스트에서 직렬화된 방식으로 덧붙이지만, 세 번째 커널의 경우는 디바이스에서 병렬화된 방식으로 padding을 구현하였다.

또한 커널의 실행 시간을 측정 할 때는, 세 커널 모두 같은 입력이 주어진 것으로 가정하여, CUDA core를 사용한 커널은 행렬의 곱셈을 계산하는 부분을, Tensor core를 사용한 커널은 padding을 덧붙이는 부분과 행렬의 곱셈을 계산하는 부분을 모두 측정하였다.

3.3 실험 결과

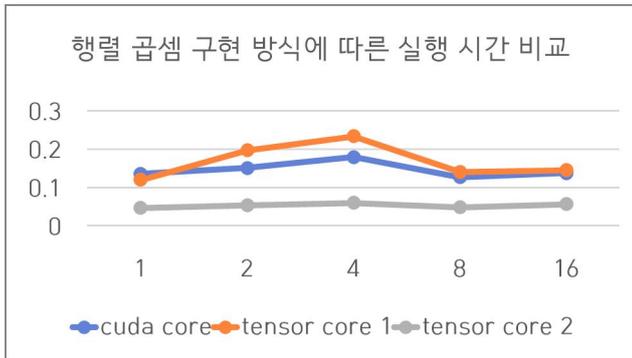


그림 1 곱셈 구현 방식에 따른 커널 실행 시간 비교

소규모 행렬 곱셈의 구현 방식에 따라 실행 시간을 비교한 결과 m=n=k=1인 경우를 제외한 모든 상황에서 실행 시간이 tensor core 1 > cuda core > tensor core 2인 것으로 나타났다.

일반적인 CUDA core보다 연산 속도가 빠르다고 알려진

Tensor core를 사용한 경우(tensor core 1)의 실행 시간이 CUDA core 보다 길게 나타난 이유는 그림 2를 통해 알 수 있는데, 이는 padding을 호스트에서 순차적인 코드로 덧붙이고 이를 디바이스의 글로벌 메모리로 옮기는 과정 때문이다.



그림 2 Nsight System을 이용하여 분석한 m=n=k=8인 소규모 행렬 곱셈에서 CUDA API의 실행 시간

반면, 소규모 행렬에 padding을 덧붙이는 과정까지 병렬화된 커널(tensor core 2)의 경우 다른 커널에 비해 두 배 이상 빠른 결과가 측정되었는데, 이는 소규모 행렬 연산에서 소규모 행렬에 padding을 덧붙여 Tensor core에서 지원하는 fragment 사이즈로 만들어주는 전처리 과정을 병렬화하는 것의 중요성을 보여준다.

4. 결론

실험 결과 Tensor core에서 지원하는 fragment의 사이즈보다 크기가 작은 소규모 행렬에서는 padding을 덧붙이는 과정을 병렬화하여 적용한 뒤, Tensor core를 사용하는 것이 더 효율적이었다. 이 전처리 과정과 데이터를 CPU에서 GPU로 옮기는 과정에 가장 시간이 많이 소모되었으며, 해당 과정을 병렬화하는 것이 성능의 결정적인 영향을 미치는 것으로 보인다.

Acknowledgement

이 연구는 2024년 정부(산업통상자원부)의 재원으로 한국 산업기술진흥원의 지원을 받아 수행된 연구임. (P0020632, 2024년 산업혁신인재성장지원사업).

참고 문헌

[1] NVIDIA Corporation. CUDA C Programming Guide. Version 12.0, 2023. [Online]. Available: <https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html>

[2] NVIDIA Corporation. Nsight Systems Get Started Guide. [Online]. Available: <https://developer.nvidia.com/nsight-systems/get-started>. [Accessed: Sep. 13, 2024].

[3] NVIDIA Corporation. CUDA Samples. [Online]. Available: <https://github.com/NVIDIA/cuda-samples>. [Accessed: Sep. 13, 2024].

[4] NVIDIA Corporation. Programming Tensor Cores in CUDA 9. [Online]. Available: <https://developer.nvidia.com/blog/programming-tensor-cores-cuda-9/>. [Accessed: Sep. 13, 2024].

혼합 정밀도 및 양자화를 활용한 객체 탐지 모델 성능 비교 연구

박준형, 김진성
 중앙대학교 컴퓨터공학과
 e-mail : acorn0415@cau.ac.kr, kimjsung@cau.ac.kr

A Comparative Study of Object Detection Model Performance Using Mixed Precision and Quantization

Junhyeong Park, Jinsung Kim
 School of Computer Science and Engineering, Chung-Ang University

요 약

최근 딥러닝 기술의 발전으로 동영상 분석 기능이 크게 향상되었으며, 이러한 기술은 주로 고성능 GPU 서버에서 많이 활용되고 있다. 그러나 최근에는 모바일 디바이스나 임베디드 시스템과 같은 저전력 환경에서도 실시간 영상 분석에 대한 수요가 급증하고 있다. 이러한 변화는 모바일 기기나 소형 디바이스에서 효율적으로 동작할 수 있는 경량화된 객체 탐지 모델의 필요성을 더욱 강조하고 있다. 기존의 고성능 모델은 높은 정확도를 제공하지만, 저전력 디바이스에서는 처리 속도나 메모리 효율성 면에서 한계를 보인다. 따라서 경량화 작업을 통해 이러한 제약을 해결하는 것이 중요하다. 본 논문에서는 혼합 정밀도 기법과 양자화 기법을 결합하여 실시간 객체 탐지에 적합한 경량화된 모델을 제안하고, 그 성능을 다양한 관점에서 평가한다. 특히, 대표적인 객체 탐지 모델인 YOLO, SSD, Faster R-CNN을 기준으로 각 모델의 성능을 비교하고, 양자화 모델, 혼합 정밀도 모델, 그리고 양자화와 혼합 정밀도 기법을 결합한 모델의 정확도와 처리 속도, 메모리 효율성을 평가하고 비교한다. 이를 통해 경량화된 모델이 기존 모델 대비 성능 저하를 최소화하며 저전력 환경에서 효율적으로 동작할 수 있는지를 확인한다.

1. 서 론

최근 딥러닝 기술은 비약적인 발전을 이루었으며, 그중에서도 컴퓨터 비전(Computer Vision, CV) 분야는 많은 응용 분야에서 큰 성과를 보였다. 이미지 분류, 객체 탐지, 동작 인식 등의 기술이 발전하면서, 다양한 산업 분야에서 딥러닝 모델을 활용한 애플리케이션이 증가하고 있다. 이와 함께 모바일 디바이스의 성능 또한 빠르게 발전하면서, 고성능 딥러닝 모델을 스마트폰, IoT 디바이스와 같은 저전력 환경에서도 실시간으로 활용하려는 수요가 커지고 있다. 그러나 딥러닝 모델은 일반적으로 높은 계산 자원과 메모리를 요구하기 때문에, 이러한 모델을 직접 모바일 디바이스에 적용하기에는 한계가 있다. 이를 해결하기 위해 최근에는 모델 경량화에 대한 연구가 활발히 진행되고 있으며, 대표적인 경량화 방법으로 혼합 정밀도(Mixed-Precision)[1] 기법과 양자화(Quantization)[2] 기법이 주목받고 있다. 본 논문에서는 이러한 경량화 기법을 활용하여, 모바일 환경에서도 실시간으로 동작 가능한 객체 탐지 모델을 제안하고, 그 성능을 다양한 지표를 통해 평가하고자 한다.

2. 관련 연구

딥러닝 모델의 경량화는 모바일 및 임베디드 환경에서

실시간 처리를 가능하게 하기 위한 중요한 연구 주제이다. 본 연구에서는 두 가지 대표적인 경량화 방법인 혼합 정밀도(Mixed-Precision) 기법과 양자화(Quantization) 기법을 활용하여 모델 성능을 개선하였다.

2.1 Mixed Precision 기법과 AMP의 활용

혼합 정밀도(Mixed-Precision)는 딥러닝 모델에서 연산 성능을 최적화하기 위해 사용된다. 이 기법에서는 계산을 부동소수점 32비트(FP32) 대신 16비트(FP16)로 수행하여 연산 속도를 높이고, 메모리 사용량을 줄인다. 중요한 값은 다시 FP32로 누적하여 모델의 정확도를 유지하면서 연산 성능을 향상시킨다. 이를 통해 메모리 사용량을 줄이고, 연산 속도를 크게 개선할 수 있다. 특히, NVIDIA의 Automatic Mixed Precision (AMP) 기법은 간단한 코드 추가로 이러한 혼합 정밀도 연산을 자동으로 적용할 수 있도록 지원하며, 이를 통해 처리 속도를 높이는 동시에 모델의 정확도를 유지하는 데 유리하다. 본 연구에서는 AMP를 사용하여 모델의 훈련에 혼합 정밀도 기법을 적용하였다.

2.2 양자화 기법

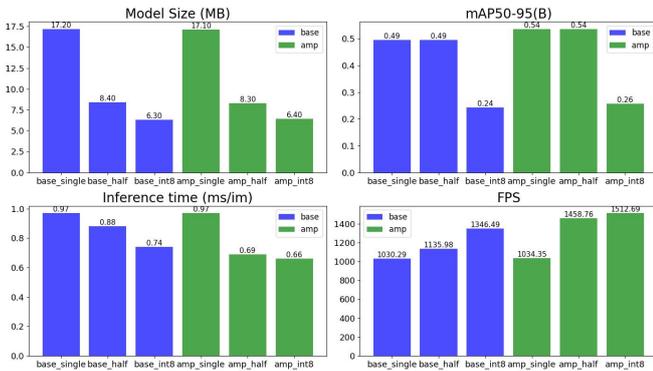
양자화는 모델의 가중치 및 활성화 값을 낮은 비트 정밀도로 변환하여 메모리 사용량을 줄이고, 연산 속도를 높이는 기법이다. 대표적으로 정수(int8) 양자화가 널리 사용되

며, 이를 통해 모델의 크기가 줄어들고, 연산 비용이 절감된다. 본 연구에서는 정적 양자화 기법을 적용하여, 사전 캘리브레이션 데이터를 활용해 가중치와 활성화 값을 양자화하였다. 정적 양자화는 입력 데이터의 분포를 기반으로 양자화 범위를 미리 결정하여, 추론 중 성능을 최적화하는 방식이다. 이를 통해 다양한 입력 데이터에서 메모리와 연산 효율성을 극대화하였다.

3. 실험

3.1 실험 환경 및 사용 모델

본 연구의 실험은 고성능 그래픽 카드인 NVIDIA RTX 4090에서 진행되었으며, 대표적인 객체 탐지 모델인 YOLO[3] 모델을 사용하여 성능을 평가하였다. YOLOv8은 최신 객체 탐지 모델로, 실시간 처리 성능과 높은 정확도를 동시에 제공한다. 본 연구에서는 Ultralytics에서 제공하는 YOLOv8 모델을 사용하여 혼합 정밀도 및 정적 양자화 기법을 적용한 후, 정확도, 처리 속도, 메모리 사용량 등의 지표를 평가하였다[4][5].



(그림 1) YOLOv8모델의 혼합 정밀도, 양자화 기법 적용 성능 비교

3.2 실험 결과

다음 그림은 YOLOv8 모델을 대상으로 AMP 적용 및 양자화(half precision, int8) 적용 유무에 따른 모델 크기 (Model Size), 정확도(mAP50-95(B)), 추론 시간(Inference Time), 초당 프레임(FPS)의 성능을 비교한 결과를 나타낸다.

- Model Size 측면에서, 양자화 기법을 적용한 경우 (base_int8, amp_int8) 모델 크기가 가장 작았으며, AMP를 적용한 half precision 모델(amp_half)은 base_half 모델에 비해 크기가 거의 차이가 없었지만 single 모델들에 비해 사이즈가 많이 줄어든 것을 볼 수 있다.
- mAP50-95(B) 성능 평가에서는 AMP를 적용한 모델들이 전반적으로 base 모델들보다 더 높은 정확도를 보였다. 특히 AMP 적용 half precision 모델(amp_half)이 가장 좋은 평가지표를 기록하였다. 반면에 int8로 양자화 된 모델들은 0.24까지 하락한 모습을 보였다.
- Inference Time 평가에서도 AMP 적용 half precision 모델(amp_half)이 양자화를 적용한 AMP모델(amp_int8)과 거의 같은 속도를 보여주었다. 또 AMP를 적용한 모델이 base모델과 차이를 보여주었다.

- FPS에서는 양자화 기법과 AMP를 같이 적용한 모델 (amp_int8)이 가장 높은 프레임 수를 기록하며 가장 빠른 처리 성능을 보였다.

결과적으로, AMP를 사용하고 half precision으로 추론한 모델이 전반적으로 가장 합리적인 성능(평가지표와 속도)을 보였으며, 모바일 환경에서도 성능 저하를 최소화하면서 실시간 객체 탐지에 적합한 경량화된 모델임을 확인하였다.

4. 결론

본 연구에서는 YOLOv8 객체 탐지 모델을 대상으로, 혼합 정밀도(Mixed Precision) 및 양자화(Quantization) 기법을 적용하여 모델 경량화를 실현하고 그 성능을 평가하였다. NVIDIA RTX 4090 환경에서 실험을 진행하였으며, AMP(Automatic Mixed Precision)와 정적 양자화 기술을 사용하여 학습 및 추론 속도, 모델 크기, 정확도를 분석하였다. 실험 결과, AMP를 적용한 모델이 전반적으로 base 모델보다 더 높은 정확도와 빠른 추론 성능을 보였다. 특히, AMP half precision 모델이 가장 우수한 정확도를 기록하였으며, int8 양자화 모델은 처리 속도에서 가장 뛰어난 결과를 나타냈다. AMP를 적용한 half precision 모델은 성능과 효율성의 균형을 가장 잘 맞춘 것으로 나타났다.

향후 연구에서는 정적 양자화 외에도 동적 양자화와 적응형 양자화를 적용하여 모델의 성능을 더욱 최적화할 계획입니다. 또한, 기존의 GPU 서버 환경에서 벗어나 모바일 및 임베디드 디바이스와 같은 저전력 환경에서의 실험을 진행하여, 이러한 기법들이 실제 모바일 환경에서도 효율적으로 작동하는지를 평가할 예정입니다.

Acknowledgement

이 연구는 2024년 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임. (P0020632, 2024년 산업혁신인재성장지원사업).

참고문헌

- [1] MICIKEVICIUS, Paulius, et al. Mixed precision training. arXiv preprint arXiv:1710.03740, 2017.
- [2] KRISHNAMOORTHY, Raghuraman. Quantizing deep convolutional networks for efficient inference: A whitepaper. arXiv preprint arXiv:1806.08342, 2018.
- [3] REDMON, J. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [4] BAGHBANBASHI, Mohammadamin; RAJI, Mohsen; GHAVAMI, Behnam. Quantizing YOLOv7: A Comprehensive Study. In: 2023 28th International Computer Conference, Computer Society of Iran (CSICC). IEEE, 2023. p. 01-05.

머신러닝을 활용한 새만금 지역의 압축지수 예측

류승연

중앙대학교 대학원 지능형에너지산업융합학과
e-mail : wooa0515@naver.com

A Study on Predicting the Compression Index of Saemangeum Region Using Machine Learning

Ryu Seung Yeon

Chung-Ang University Department of Intelligent Energy and Industry

요 약

본 연구에서는 머신러닝 기법을 이용하여 새만금 지역의 압축지수를 예측하였다. 총 4가지 머신러닝 모델이 사용되었고 R^2 과 RMSE로 모델의 성능을 평가하였다. 새만금 지역을 3개 구역으로 나누어 예측한 결과 구역마다 예측 정확도에 비교적 큰 차이가 있었으며 이를 통해 예측 시 지형 및 지질학적 요인을 종합적으로 고려해야 한다는 점을 확인할 수 있었다.

유한요소해석을 이용한 말뚝의 주면마찰력에 영향을 미치는 매개변수연구

이태겸
중앙대학교 토목공학과 석사과정
e-mail : gyeom2ya0207@naver.com

A Study on the parameters affecting the friction force of the surface of a pile using finite element analysis

Lee Tae Gyeom
Graduate Student, School of Civil Engineering, Chung-Ang University

요 약

본 연구에서는 대심도 연약지반에서 마찰말뚝의 효율성을 확인하기 위하여 마찰말뚝의 마찰 지지력에 영향을 미치는 매개변수를 유한요소해석 프로그램인 PLAXIS 2D를 활용하여 연구를 진행하였다. 해석 결과, 말뚝의 극한 주면마찰력의 공식과 같이 말뚝에 인접한 지반의 점착력과 내부마찰각에 따라 말뚝의 마찰 지지력이 상이한 것으로 나타났다.

제스처로 유발된 표면의 진동을 기반으로 한 허공 제스처 인식 연구

김민혁, 김효수

중앙대학교 컴퓨터공학과

e-mail : alsgur7859@cau.ac.kr, hskimhello@cau.ac.kr

Surface-Assited In-Air Gesture Recognition

Minhyeok Kim, Hyosu Kim

School of Computer Science and Engineering, Chung-Ang University

요 약

허공에서의 손 제스처 인식은 단순하고 직관적이고 자연스러운 제스처를 통해 스마트홈, 원격 미팅 중 발표, 영상 시청 등 많은 상황에서 사용자가 편리하게 입력하고 하는 일에 몰입할 수 있도록 돕는다. 이때 상용 기기인 스마트폰과 스마트워치만 가지고 이것이 가능하다면 보편적인 사용성과 대안적인 비접촉식 인터랙션 방식의 기능까지 갖출 수 있다. 본 논문에서는 스마트폰과 스마트워치만을 이용하여 테이블 상에서 이루어지는 허공 손 제스처 인식을 가능하게 하는 SURFING을 제안한다. 손 제스처 인식을 위해 사용자는 두 팔꿈치를 테이블 위에 두고 한쪽 팔꿈치 옆에는 스마트폰을, 왼손에는 스마트워치를 차고 있다. 이때 손 제스처를 취하면 발생하는 진동이 팔, 팔꿈치를 지나 테이블을 타고 스마트폰에, 그리고 스마트워치에 도달하면 그 진동에 대한 스마트 기기의 IMU 센서 데이터를 가지고 손 제스처를 감지하고 분류한다. 구체적으로는 스마트폰의 IMU 센서 데이터에서 추출한 feature들을 이용해 1차로 손 제스처의 classification을 진행하고, 2차로 스마트워치의 IMU 센서 데이터에서 추출한 feature들을 이용해 손 제스처 classification을 보정하여 정확도를 높인다. 우리는 바뀔 수 있는 다양한 환경, 예를 들면 테이블의 종류, 기기의 종류, 스마트폰과 팔의 위치, 책상위의 장애물 등에 변화를 주며 손 제스처 인식의 정확도를 테스트 했고, 그 결과 다양한 환경에서 SURFING이 높은 정확도를 보이는 것을 확인했다.

대형 언어 모델에서 가지치기와 혼합 정밀도 기법을 활용한 성능 비교 연구

박준형, 김진성

중앙대학교 소프트웨어학과

e-mail : acorn0415@cau.ac.kr, kimjsung@cau.ac.kr

A Comparative Study on the Performance Optimization of Large Language Models Using Pruning and Mixed Precision Techniques

Park Junhyeong, Kim Jinsung

School of Computer Science and Engineering, Chung-Ang University

요 약

최근 몇 년간 자연어 처리(NLP) 분야에서 심층 학습 기술의 급격한 발전으로 인해 대형 언어 모델(Large Language Model, LLM)이 등장하게 되었다. 이러한 대형 언어 모델들은 번역, 텍스트 생성, 요약 등 다양한 작업에서 뛰어난 성능을 발휘해왔으나, 그 성능을 유지하기 위해 필요한 방대한 파라미터 수와 계산 자원으로 인해 훈련 및 배포 과정에서 높은 운영 비용이 발생하는 문제가 있다. 이를 해결하기 위한 방법으로 가지치기(pruning)와 혼합 정밀도 계산(mixed-precision computation) 같은 최적화 기법들이 제안되었다. 가지치기는 모델 파라미터의 수를 줄여 메모리 사용량과 연산량을 감소시키는 기법으로, 상대적으로 덜 중요한 파라미터를 제거해 모델의 효율성을 높인다. 반면, 혼합 정밀도 계산은 단일 정밀도(floating point 32)와 절반 정밀도(floating point 16)를 적절히 조합하여 메모리 사용량을 줄이고, 계산 속도를 향상시킨다. 본 연구는 가지치기로 인해 발생한 다양한 희소성(sparsity) 수준에서 혼합 정밀도 계산 기법이 대형 언어 모델 성능에 미치는 영향을 평가한다.

하드웨어와 소프트웨어 관점에서 본 GPU 에너지 소비량 감축 방법에 대한 연구

김인서, 김진성

중앙대학교 컴퓨터공학과

e-mail : inseo764@cau.ac.kr, kimjsung@cau.ac.kr

A Study on Hardware-Centric and Software-Centric Approaches to Reduce Energy Consumption in GPUs

Inseo Kim, Jinsung Kim

School of Computer Science and Engineering, Chung-Ang Univ.

요 약

환경 문제의 중요성이 갈수록 강조되고 있는 상황에서, 프로그램과 컴퓨팅 자원의 평가 지표에도 변화가 찾아오고 있다. 현대 사회는 기후 변화로 인한 다양한 문제를 직면하고 있다. 게다가 상당한 연산 능력이 필요한 응용프로그램의 등장도 점점 더 찾아지면서, 연산 가속기로서 GPU를 사용하는 경우가 늘어나고 있다. 이런 상황에서, 환경 문제와 기업들의 경제적 비용으로 이어질 수도 있는, 에너지 소비와 효율성은 GPU 사용에 있어서 새로운 평가 기준으로 떠오르고 있다. 따라서, 이 서베이에서는 GPU를 사용할 때 에너지 소비를 줄이기 위한 연구들을 살펴본다. 또한 그 결과를 하드웨어 중심의 관점과 소프트웨어 중심의 관점으로 분류하였다. 하드웨어 중심의 관점에서는 DVFS(동적 전압 및 주파수 조정)가 있고, GPU의 에너지 사용량을 정확하게 측정할 수 있는 도구가 제안되기도 했다. 소프트웨어 중심의 관점에서는 다양한 소프트웨어 특성이 GPU의 전력 및 에너지 소비에 미치는 영향을 알아보았다. 프레임워크와 플랫폼 조합에 따라서 에너지 효율성이 달라졌으며, 데이터 타입 또한 영향을 미쳤다. 그리고 CNN 네트워크의 레이어를 4가지로 분류하여 에너지 소비를 분석하였고, 커널을 구성할 때, 스트레드블록 크기에 따라서도 에너지 효율성에 차이가 있었다.

합성데이터 기반 소나무재선충병 탐지

정용훈*, 서상현**

중앙대학교 예술공학대학원 응용예술공학과*

중앙대학교 예술공학부**

e-mail : dydgns2017@cau.ac.kr, sanghyun@cau.ac.kr

Pine wilt disease detection based on synthetic data

Yonghoon Jung*, Sanghyun Seo**

Department of Applied Art and Technology, Chung-Ang University*

School of Art and Technology, Chung-Ang University**

요 약

소나무재선충병(Pine Wilt Disease, PWD)은 치료법이 없어 산림 관리에 심각한 문제를 일으키고 있다. 최근에는 딥러닝 기반의 모델로 PWD를 탐지하려는 연구가 진행되었으나, 대규모 고품질 데이터셋 확보는 여전히 도전 과제이다. 본 논문에서는 3D 렌더링으로 생성한 합성데이터와 딥러닝 객체 탐지 기술을 결합하여 PWD 탐지 파이프라인을 제안한다.

델타 맵 기반 고충실도 비디오 얼굴 재노화

김범수*, 서상현**

중앙대학교 일반대학원 응용예술공학과*

중앙대학교 예술공학부**

e-mail : bumsookim@cau.ac.kr, sanghyun@cau.ac.kr

High fidelity video facial re-aging based on delta map

Bumsoo Kim*, Sanghyun Seo**

Department of Applied Art and Technology, Chung-Ang University*

School of Art and Technology, Chung-Ang University**

요 약

본 논문에서는 입력된 비디오에서 떨림 없는 자연스러운 얼굴 재노화 기술을 제안한다. 기존의 한 장의 이미지에서 진행하는 얼굴 재노화 모델과 다르게, 우리의 모델은 비디오 레벨에서 얼굴 재노화를 진행할 수 있는 최초의 논문이다. 입력된 사람의 Identity 보존을 위해서 이미지 전체를 생성하는 대신, 재노화에 해당하는 델타 맵을 생성한다. 이는 주름과 같은 세부 얼굴 속성을 Identity 저하없이 쉽게 제거/생성해낼 수 있다. 우리의 모델은 여러 입력 조건에서도 원활하게 편집가능하도록 최근 생성 모델 유형인 Diffusion 모델을 기반으로 하고 있으며, 연산 효율성을 위해서 이미지 기반 편집 Diffusion 모델의 결과물을 광학 흐름으로 변형하고 사용한다. 실험에서는 다른 비디오 기반 편집 방법론 보다 자연스러운 편집 결과를 보여주었으며, 우리의 모델이 입력 얼굴의 Identity를 가장 잘 보존하는 정성적 결과를 보여준다. 추가적으로 편집의 유연함을 위해서 프롬프트 기반의 이미지 편집 기술을 채택하여 사용자가 얼굴 영역을 지정해주거나 추가적인 얼굴 탐지 모델을 사용하지 않고 효율적으로 비디오 얼굴 재노화를 할 수 있다.

합성데이터를 활용한 실내 시설 내 방문자 탐지

신원섭*, 강윤제**, 서상현**
중앙대학교 첨단영상대학원 영상학과*
중앙대학교 예술공학부**
e-mail : wonseop218@cau.ac.kr

Visitor detection in indoor facilities using synthetic data

Wonseop Shin*, Yoonje Kang**, Sanghyun Seo**
Graduated School of Advanced Imaging Science, Multimedia & Film,
Chung-Ang University*
School of Art and Technology, Chung-Ang University**

요 약

본 연구는 실내 밀집 시설에서의 방문자 탐지를 위한 합성데이터 기반 딥러닝 모델 학습을 제안한다. 실제 환경에서의 데이터 수집은 개인정보 보호 문제 및 접근 제한으로 인해 어려움이 있으며, 데이터 주석 작업 또한 많은 시간과 자원이 소요된다. 이를 해결하기 위해 본 연구는 렌더링 프로그램을 통해 실내 밀집 시설 환경을 생성하고, 인체 모델을 활용한 합성데이터를 구축하였다. 합성된 데이터셋을 바탕으로 방문자 탐지 모델을 학습하고, 실제 방문자 데이터를 통해 모델 성능을 평가하였다. 그 결과, 방문자 탐지 모델은 정밀도 96.3%, 재현율 89.2%, F1-스코어 92.4%를 기록하며, 합성데이터 기반 학습이 실제 데이터에서도 유의미한 성능을 보였음을 확인하였다.

서버리스 컴퓨팅을 위한 워크로드 예측 기술 동향 조사

조민규¹, 이재환², 박상오^{3*}
중앙대학교 대학원 컴퓨터공학과¹
국립공주대학교 컴퓨터공학과²
중앙대학교 소프트웨어학부^{3*}

e-mail : mgjo@cslab.cau.ac.kr¹, jhnlee@kongju.ac.kr², sopark@cau.ac.kr^{3*}

A Survey on Research Trends in Workload Prediction for Serverless Computing

Mingyu Jo¹, Jaehwan Lee, Sangoh Park

Department of Computer Science and Engineering, Graduate School,
Chung-Ang University¹

Department of Computer Science and Engineering, Kongju National
University²

School of Computer Science and Engineering, Chung-Ang University^{3*}

요 약

본 논문에서는 서버리스 컴퓨팅의 최신 연구 동향을 제시한다. 특히 서버리스 컴퓨팅의 주요 이슈 중 하나인 콜드 스타트(Cold Start) 문제를 해결하기 위해 적용되는 다양한 머신 러닝 기반의 워크로드 예측 기법들은 분석하고, 그 트렌드를 분석한다. 또한 본 연구에서는 워크로드 예측 기법들을 개발하기 위한 실세계 워크로드 데이터셋의 중요성을 분석한다.

1. 서 론

서버리스 컴퓨팅 [1] 은 물리적인 서버 구축과 배치를 고려 할 필요 없이 기능의 실행 단위로 리소스를 활용할 수 있는 차세대 클라우드 컴퓨팅 패러다임이다. 서버리스 컴퓨팅의 기능 실행 단위는 주로 람다(Lambda) 함수로 불리며, 이 독립 실행 단위인 람다 함수가 호출 될 때 클라우드 리소스가 동적으로 할당되었다가 실행이 완료되면 해제되는 방식으로 동작하기 때문에 서비스 제공자는 리소스 활용 효율을 크게 개선할 수 있다.

본 논문에서는 서버리스 컴퓨팅을 달성하기 위한 주요 해결 과제 중 하나인 콜드 스타트 문제 해결을 위한 최근의 접근 방법론을 탐색한다. 특히 최근 연구는 머신 러닝 기법을 적용하여 서버리스 워크로드의 예측을 시도하고 있다. 이에 본 논문의 2장에서는 최근 연구들을 소개 및 비교 분석한 뒤, 3장에서 결론을 맺는다.

2. 관련 연구

서버리스 컴퓨팅의 독립 실행 단위인 람다 함수는 상태를 유지하지 않는다. 람다 함수의 실행 요청은 리소스 할당을 통한 런타임 환경 준비, 함수 실행, 그리고 리소스 해제로 이어진다. 람다 함수의 실행은 온 디맨드 방식이므로 런타임 인스턴스를 동적으로 할당해야 하며, 이는 함수를 반복적으로 호출하는 시나리오에서 비효율성을 초래할 수 있다. 람다 함수를 호출할 때마다 새로운 람다 인스턴스와 런타임을 필요로 하는데, 이 프로비저닝 과정을 콜드

스타트(Cold Start) 라고 부른다. 콜드 스타트는 서버리스 워크로드의 전체 실행 성능(응답 속도, 실행 시간 등)에 상당한 영향을 미칠 수 있기 때문에 이러한 문제를 해결하기 위해 람다 인스턴스의 캐싱이나 사전 프로비저닝에 대한 연구들이 수행되고 있다.

최근의 연구들은 콜드 스타트 문제를 완화하기 위해 서버리스 워크로드를 보다 정확하게 예측하는 것을 목표로 한다. Pan 등은 [2] 워크로드 분석을 통해 람다 함수의 호출량의 증가·감소 추세를 그래디언트로 측정하였다. 이러한 추세에 따라 인스턴스를 사전 할당하거나 리소스를 회수한다. Noble 등은 [3] 머신 러닝 기반 서버리스 워크로드 예측 기법이 효과적임을 실험적으로 입증하였다. Long Short-Term Memory(LSTM), Autoregressive Integrated Moving Average (ARIMA), Vector Autoregression(VAR) 등의 시계열 예측 모델을 설계하여 서버리스 워크로드의 예측 성능을 평가한 실험에서는 LSTM 모델이 가장 효과적임을 보고하였다. Golec 등은 [4] 람다 인스턴스의 콜드 스타트 지연시간을 머신 러닝 기법을 이용해 예측하였다. 선형회귀, LSTM 등 7개의 머신 러닝 모델에 대한 콜드 스타트 지연시간 예측 성능을 평가한 실험에서는 XGBoost 모델이 가장 좋은 성능을 보인다고 보고하였다.

3. 결론 및 고찰

서버리스 컴퓨팅에서 콜드 스타트 문제는 여전히 개선

되어야 할 주요 문제 중 하나이며, 이 분야에서 다양한 연구들이 수행되고 있다. 최근 트렌드는 머신 러닝을 사용하여 램다 인스턴스의 워크로드나 대기 시간을 예측한 뒤 이를 인스턴스 프로비저닝에 활용하는 것으로 분석된다. 이러한 연구들은 머신 러닝 모델의 예측 정확도 향상이 콜드 스타트로 인한 지연 시간 감소로 이어질 수 있음을 보고하고 있다. 그러나 표 1에 기술한 바와 같이 머신 러닝 기반의 서버리스 컴퓨팅 예측 연구는 아직 초기 단계에 있으며, 개별 연구마다 사용된 데이터셋과 워크로드 생성 방식이 상이하다. 향후 연구에서는 서버리스 워크로드를 대표할 만한 데이터셋의 생성과 공통된 평가 지표를 통해 다양한 서버리스 워크로드 예측 모델의 성능을 비교할 수 있는 공통된 벤치마크 프레임워크의 구축이 필요하다.

(표 1) 기존 연구들의 데이터셋 및 워크로드 생성 방식의 비교

연구	데이터셋	시뮬레이션 프레임워크
[2]	Azure Functions Trace	OpenWhisk
[3]	Azure Dataset	X
[4]	Predictive Maintenance Dataset	JMeter, Google Cloud Functions

Optimization and Automatic Workflow Resource Scheduling," IEEE Transactions on Sustainable Computing, Vol. 40 No. 3, March 2023, pp. 1-12.

[3] Noble N. T., Dev Y. P., Joseph C. T., "Machine Learning Based Techniques for Workload Prediction in Serverless Environments," 2023 International Conference on Electrical, Electronics, Communication and Computers (ELEXCOM), Roorkee, India, 2023, pp. 1-6.

[4] Golec M., et al., "MASTER: Machine Learning-based Cold Start Latency Prediction Framework in Serverless Edge Computing Environments for Industry 4.0," IEEE Journal of Selected Areas in Sensors, Vol. 1 2024, pp. 36-48.

Acknowledgement

본 연구는 2024년 과학기술정통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2024-0-00073)

참고 문헌

[1] 김동민, 손재기. "서버리스 컴퓨팅 기술 동향", 『한국통신학회지(정보와통신)』 제37권 제8호, 2020. pp.39-45.

[2] Pan S., Zhao H., Cai Z., Li D., Ma R., Guan H., "Sustainable Serverless Computing with Cold-start

지능형 문서처리 플랫폼의 중요정보 보호 방안 연구

홍기완¹, 한소영¹, 김자원², 김양훈³, 장항배^{4*}

중앙대학교 일반대학원 융합보안학과¹

대진대학교 스마트융합보안학과²

신한대학교 사이버드론봇군사학과³

중앙대학교 산업보안학과^{4*}

e-mail : {ghd9201¹, hsy0530¹, hbchang^{4*}}@cau.ac.kr, jjawon@daejin.ac.kr²,
kimyh7902@shinhan.ac.kr³

A Study on Protecting Sensitive Information of Intelligent Document Processing Platform

Giwan Hong¹, Soyoung Han¹, Jawon Kim², Yanghoon Kim³,
Hangbae Chang^{4*}

Dept. of Security Convergence, Graduate School, Chung-Ang University¹,

Dept. of Smart Convergence Security, DaeJin University²,

Dept. of Cyber Drone Bot Military Engineering, Shinhan University³,

Dept. of Industrial Security, Chung-Ang University^{4*}

요 약

오늘날 디지털 업무환경에서 기업과 조직의 문서자산이 급격히 증가함에 따라, 업무의 생산성과 효율성 향상을 위한 지능형 문서 처리 플랫폼 도입이 중요한 과제로 부상하고 있다. 기업의 문서자산에는 영업비밀, 핵심기술 정보 등 중요정보가 포함될 수 있으며, 중요정보를 효과적으로 탐지하고 정보의 가치에 따른 차등 분류와 보안관리를 통해 데이터의 안전한 활용을 보장하는 것이 필수적이다. 중요정보 탐지와 문서자산의 보안 분류를 위한 다양한 방안이 존재하지만, 일반적으로 문서의 정보를 효과적으로 탐지하고 내재된 가치를 파악하거나 분류의 일관성을 확보하기 어려운 한계가 있다. 이에 본 연구는 자연어 처리(Natural Language Processing, NLP), 지능형 문서 처리(Intelligent Document Processing, IDP), 메타데이터 및 데이터 보호 관련 기술들을 분석하고 지능형 문서 처리 플랫폼에 효과적인 중요정보 탐지 및 보호 방안을 검토하고자 한다. 이를 통해 기업과 조직이 중요정보를 안전하게 보호하면서도 문서자산을 효과적으로 활용할 수 있도록 하는 지능형 문서 처리 플랫폼 도입을 위한 기초자료를 제공하고자 하며, 기업의 데이터 보안 수준을 강화하는 동시에 업무 생산성을 향상시키는 데 중요한 역할을 할 것으로 기대된다.

감사의 글

“이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2021-II210116, 문서 데이터의 정보화 및 협업을 위한 지능형 문서 처리 플랫폼 기술)”

탐색적 데이터 분석(EDA)기반 흡연상태와 음주여부에 따른 BMI수치 분석

전석원
 한남대학교 컴퓨터공학과
 e-mail : presium575730@gmail.com

장준혁
 한남대학교 컴퓨터공학과
 e-mail : jhjang@hnu.kr

BMI level analysis according to smoking status and drinking status based on exploratory data analysis (EDA)

Seokwon Jeon
 Department of Computer Engineering
 Hannam University

Joonhyouk Jang
 Department of Computer Engineering
 Hannam University

요약

본 논문은 국민건강보험공단에서 제공한 2021년과 2022년의 건강검진 데이터를 활용하여, 흡연과 음주가 체질량 지수(BMI)에 미치는 영향을 분석하였다. 이를 위해 탐색적 데이터 분석(EDA)을 기반으로 흡연 여부와 음주 여부에 따른 BMI의 변화를 비교하고, 두 요인 간의 상관관계를 조사하였다. 또한, 본 EDA과정을 통해 흡연과 음주가 BMI에 미치는 영향이 어느 정도인지 결론을 도출하였으며, 향후 관련 연구 진행 시 보완되어야 할 점을 제시하였다.

1. 서론

현대 사회에서 건강 관리에 대한 관심과 중요성은 점점 더 커지고 있다. 관련 지식들을 쉽게 얻을 수 있는 환경 속에서, 우리의 건강을 위협하는 주요 요인으로 손꼽히는 것이 바로 흡연과 음주이다. [1][2]에 따르면 흡연과 음주는 심혈관 질환, 당뇨병, 암 등 여러 질병의 발병 위험을 증가시킨다. 그러므로 흡연상태와 음주여부가 개인의 건강 상태에 어떤 영향을 미치는지 알게 되는 것은 효과적인 건강 관리에 큰 도움이 될 것이다.

본 논문에서는 흡연과 음주가 건강에 미치는 영향을 분석하기 위하여 BMI(체질량지수)를 주요 지표로 사용하였으며, 데이터 분석 기법인 탐색적 데이터 분석(EDA)을 기반으로 연구를 진행하였다. 본 연구에서 활용한 데이터는 국민건강보험공단에서 제공하는 21년도와 22년도 건강검진정보 데이터이다. 건강검진정보는 해당년도에 건강검진을 수검한 국민건강보험 가입자 100만명에 관한 기본정보와 검진내역으로 구성되어있으며, 20세이상을 대상으로 하였다.

2. EDA 기반 데이터 분석

EDA는 데이터를 분석하고 결과를 도출하는 기법 중 하나로, 이 기법의 목표는 통계적인 가설 검증이 아닌 데이터의 탐색과 인사이트 도출에 있다. EDA를 수행하면서 데이터의 결측치와 이상치를 처리하고 데이터를 시각화하는 과정을 거치게 된다[3].

2.1 데이터 정제

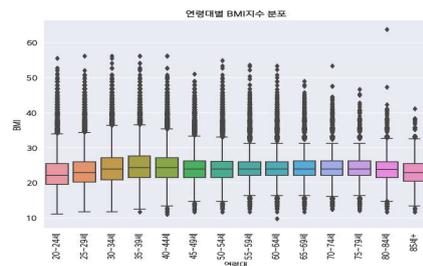
데이터 분석을 하기에 앞서, 건강검진정보 데이터에 대해 정제를 진행하였다. 먼저, 숫자 코드로 되어 있던 ‘성별’, ‘나이’, ‘지역’, ‘흡연상태’, ‘음주여부’ 열을 문자열로 변환하였다. 이어서 결측치를 확인하였는데, 200만 행의 데이터 중 분석에서 사용하는 ‘흡연상태’와 ‘음주여부’ 열에서 각각 433, 244개의 결측치가 존재하였다. 분석의 정확성을 위하여 다른 값으로 대체 하지 않고 해당 행을 제거하였다. 또한, 본 데이터에는 BMI열이 존재하지 않아 신장과 체중을 이용하여 BMI열과 비만분류 열을 추가하였다[4].

기준 년도	지역 성별	연령대	신 장	체 중	허리 둘레	최고혈 압	최저 혈압	식전혈당 (공복혈 당)	혈청지오 티(AST)	혈청지오 티(ALT)	간염지 티	흡연 상태	음주여 부	BMI	비만 분류
0	2022	여 자	65-69 세	160	75	89.0	127.0	73.0	84.0	25.0	20.0	15.0	비흡 연자 입니다	29.3	1단계 비만
1	2022	여 자	40-44 세	150	45	68.0	127.0	75.0	85.0	19.0	10.0	7.0	비흡 연자 입니다	20.0	정상

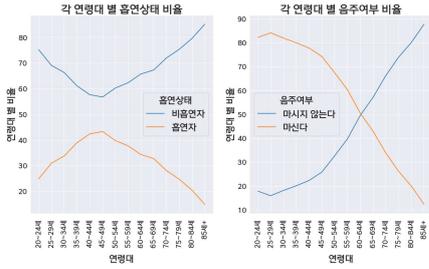
(그림 1) 정제 완료한 데이터

2.2 데이터 경향 확인

음주와 흡연에 따른 BMI 분석을 위하여 데이터의 전체적인 경향을 그래프로 시각화하였다.



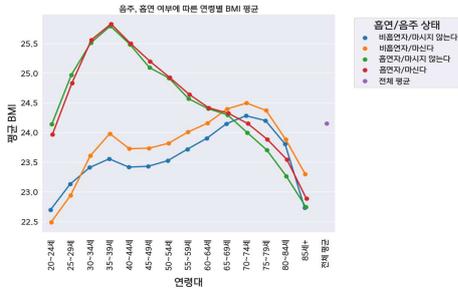
(그림 2) 연령대별 BMI지수 분포



(그림 3) 각 연령대 별 흡연과 음주 여부 비율

먼저, 그림 2에서 연령대별로 박스의 크기와 평균이 비교적 균일하므로 데이터가 균등하게 분포되어있는 것을 확인할 수 있다. 다음으로 그림 3의 흡연상태 비율을 보면 데이터에서 전체적으로 흡연자의 비율이 낮은 것으로 나타나고, 음주여부 비율에서는 연령대가 높아짐에 따라 음주자의 비율이 감소하는 경향이 나타나고 있다.

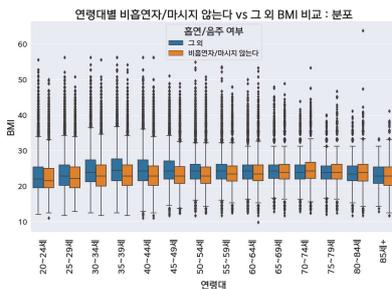
2.3 흡연 및 음주가 BMI에 미치는 영향



(그림 4) 음주, 흡연여부에 따른 BMI 평균

2.2에서 확인한 경향을 바탕으로 본 주제에 관한 비교를 하기 위해 그림 4와 같이 시각화를 하였다. 그림 4를 보면 음주와 흡연 모두 하는 사람과 흡연만 하는 사람의 BMI 평균의 차이가 적고, 흡연만 하는 사람의 BMI가 음주만 하는 사람의 BMI보다 높게 나타나는 것을 알 수 있다.

연령대로 보았을 때 40세에서 44세까지는 비슷한 경향을 보이다가 60세에서 69세부터 '비흡연자/마신다'와 '비흡연자/마시지 않는다'가 '흡연자/마신다'와 '흡연자/마시지 않는다'의 BMI를 추월하는 경향을 보인다. 70세부터는 BMI가 떨어지며, 85세+의 경우 다른 연령에 비해 압도적으로 수치가 낮게 나타난다.



(그림 5) 비흡연 및 비음주자와 그 외의 BMI 비교

흡연 및 음주와 BMI의 관계를 더욱 정확히 파악하기 위해, 흡연과 음주 전부 하지 않는 데이터와 그 외의 데이터를 비교하는 박스 그래프인 그림 5를 시각화하였다.

그림 5의 30세에서 59세까지는 흡연자와 음주자 중 어느 하나 이상 해당하는 '그 외' 그룹의 BMI가 높게 나타나고 있으나, 연령대가 높아지면서 점점 차이가 줄어들어가는 경향을 보인다. 65세부터는 흡연과 음주를 하지 않는 그룹의 BMI가 '그 외' 그룹의 BMI보다 높게 나타난다. 이러한 경향으로 보아 음주와 흡연 외의 요인이 작용한 것으로 보인다.

3. 결론 및 향후 보완점

본 논문에서는 흡연상태와 음주여부가 BMI에 어떤 영향을 주는지 분석하였다. 그 결과, 흡연과 음주가 BMI에 일정 부분 영향을 미치는 것은 부정할 수 없으나, 연령대에 따라 그 수치가 다르게 나타나는 것을 확인할 수 있었다.

본 연구에서 사용된 데이터와 분석 기법을 통하여 어느 정도의 분석은 가능하였으나, BMI에 영향을 미칠 수 있는 다른 요인(운동, 식습관, 유전적 요인 등)이 고려되지 않았고[5], 관련 데이터가 부족하여 분석에 한계가 존재했다.

따라서 향후 연구에서는 흡연과 음주 이외에도 위에서 서술한 다른 건강 관련 요인들을 포함한 다변량 분석을 통하여 종합적으로 영향을 분석할 필요가 있을 것이다. 또한, 연령대별로 세분화하여 비교하는 방법이나 더 나아가서 인구통계학적인 변수를 추가로 고려한다면 더욱 정확한 결론이 나올 것으로 기대된다.

참고문헌

[1] Banks E, Joshy G, Korda R.J, Stavreski B, Soga K, Egger S, Day C, Clarke N.E, Lewington S, Lopez A.D, "Tobacco smoking and risk of 36 cardiovascular disease subtypes: fatal and non-fatal outcomes in a large prospective Australian study", BMC Medicine, Vol.17 No.1, July 2019. pp.128.

[2] Zhang X, Liu Y, Li S, Lichtenstein A.H, Chen S, Na M, Veldheer S, Xing A, Wang Y, Wu S, Gao X, "Alcohol consumption and risk of cardiovascular disease, cancer and mortality: a prospective cohort study", Nutrition Journal, Vol.20, No.1, February 2021, pp.13.

[3] 홍명우, 강윤희. "환경 모니터링을 위한 EDA 기반 데이터 분석", 『디지털콘텐츠학회논문지』, 제23권 제7호, 2022. pp.1289-1295.

[4] 고성훈. "체질량지수의 비만도 분류와 주관적 체형인식의 일치, 불일치에 따른 삶의 질 및 정신건강 관련성", 『고려대학교 국내석사학위논문』, 2018.

[5] 이현주. "현역병의 건강행태, 주관적 체형인식, 체중조절 활동이 BMI 변화에 미치는 영향", 『한국산학기술학회논문지』, 제16권 제5호, 2015. pp.3353-3360.

비지도 기계학습을 이용한 뇌 영상 기반 자폐 스펙트럼 장애의 생물학적 특성 분석기술 개발¹⁾

최충민^{1,†}, 김상우^{1,†}, 김예은², 조한규^{1,2,*}

1. 국립군산대학교 컴퓨터소프트웨어특성화대학 인공지능융합학과

2. 국립군산대학교 전자정보공학부 정보통신전파공학전공

e-mail: thdlthddl2@gmail.com, gnbupi@naver.com,

kimye0729@kunsan.ac.kr, ljwjw18@kunsan.ac.kr, hgjo@kunsan.ac.kr*

Biological Characteristics of Autism Spectrum Disorder: Brain Image Analysis Using Unsupervised Machine Learning

Chungmin Choi¹, Sangwoo Kim¹, Ye-eun Kim², Han-Gue Jo^{1,2,*}

1. Department of AI Convergence, College of Computer and Software, Kunsan National Univeristy

2. School of Electronic and Information Engineering, Kunsan National Univeristy

요 약

자폐 스펙트럼 장애(Autism Spectrum Disorder, ASD)는 개인마다 다양한 증상과 행동 양식을 나타낸다. 표면적인 증상을 기반으로 한 ASD의 진단 및 분류 방법은 주관적인 요소가 개입될 수 있으며, 정확한 진단과 그에 따른 적절한 처방에 한계가 있다. 이에 본 연구는 기존의 증상 중심 접근법을 넘어, 데이터 기반 연구 방법을 통해 뇌 영상을 기반으로 ASD 환자의 생물학적 특성을 도출하고자 한다. 그래프 신경망(Graph Convolutional Network, GCN)과 오토인코더(Auto Encoder)를 결합하여 ASD 환자의 복잡한 뇌 네트워크 특성을 학습하고, 학습된 모델을 해석하여 ASD 환자의 주요 뇌 특징을 도출하였다. 그 결과, 기저핵(Basal Ganglia)과 배측 전두엽(Dorsal Frontal Cortex) 연결성, 방추회(Fusiform Gyrus)와 측소뇌(Lateral Cerebellum) 연결성, 그리고 전방 전두피질(Anterior Prefrontal Cortex)과 하두정소엽(Inferior Parietal Lobule) 연결성을 도출하였다. 향후 다양한 정신질환 환자와 비교 분석을 통해 도출된 뇌 연결성이 ASD 환자만의 특성인지를 검증할 필요가 있다. 이러한 연구결과는 기존의 증상 중심 접근법을 넘어, 생물학적 데이터를 기반으로 한 새로운 ASD 환자의 특성을 도출하고 더 나아가 ASD의 생물학적 아형을 분류하는 데 기여할 것으로 기대된다.

1. 서 론

자폐 스펙트럼 장애(Autism Spectrum Disorder, ASD)는 개인마다 다양한 증상과 행동 양식을 보인다. 기존 증상 중심의 진단 방법은 겉으로 표현되는 증상에 기반하여 생물학적 특성과 다를 수 있고 주관적 요소가 개입될 수 있어, 정확한 진단과 처방에 한계가 있다[1].

최근 인공지능 및 기계학습 기법의 급속한 발전으로, 이를 활용하여 임상 데이터와 뇌 영상 데이터를 분석하는 연구가 활발하게 진행되고 있다[2]. 증상에 기반하여 분류된 ASD 환자를 정상인과 구분할 수 있는 인공지능 모델들이 개발되었으며, 그 정확도도 크게 향상되었다. 하지만, 데이터를 기반으로 ASD 환자의 뇌 생물학적 특성을 도출하는 연구는 찾아보기 힘들다.

본 논문에서는 기존의 증상 기반의 접근법이 아닌 대규모 뇌영상 데이터를 기반으로 비지도학습 기법을 이용하여 뇌 생물학적 특성에 기초한 ASD 환자의 뇌 특성을 도출하고자 한다.

2. 방법

2.1 데이터 수집 및 전처리

본 연구에서는 여러 기관에서 측정된 394명의 대규모 뇌영상 데이터(기능적 자기공명영상, fMRI)를 활용하여 [3], ASD 환자의 뇌 생물학적 특성을 도출한다. 휴지기 상태에서 뇌의 활성화를 간접적으로 측정하는 Blood Oxygen Level Dependent(BOLD) 신호를 통해 fMRI 데이터를 수집하였다. fMRI 데이터는 Dosenbach's 뇌 템플릿을 이용하여 160개 뇌 영역으로 세분화하고, 각 영역의 BOLD 시계열을 추출, 뇌 영역 간의 기능적 연결성(총 연결성 $160 \times 159 / 2 = 12,720$)을 상관 계수로 수치화하여 모델의 입력값(인접 행렬)으로 사용하였다. 상관 계수의 변동 폭을 확장하기 위해 Fisher-z 변환을 사용하여 정규화하였다.

2.2 GCN (Graph Convolutional network)

뇌의 연결성 데이터가 가지는 복잡한 네트워크 구조를 효과적으로 분석하기 위해 그래프 콘볼루션 신경망(GCN)을 적용하였다. 이는 분할된 뇌 영역 간의 연결성을 학습하고, 이를 기반으로 ASD 환자의 뇌 네트워크를 분석할 수 있도록 설계하였다.

GCN은 그래프 구조 데이터(예를 들어, 인접 행렬)에서 동작하도록 설계된 심층 학습 모델로 복잡한 네트워크에

1) 본 연구는 국립대학육성사업비, 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065)

† These authors contributed equally to this work

서 효율적인 표현 학습과 관계 및 종속성을 분석할 수 있다. 노드와 엣지로 구성된 그래프에서 노드 간의 관계를 모델링하고 각 노드의 특성 및 전체 그래프 구조를 학습한다. GCN은 그래프 데이터에 대해 컨볼루션 연산을 수행하도록 설계된 네트워크이며, 그래프의 구조적 특성을 반영하여 인접 노드 간의 정보 교환을 효율적으로 처리한다.

2.3 오토인코더(Auto Encoder, AE) 기반 모델 학습

입력된 인접 행렬로부터 주요 뇌 특징을 추출하기 위해 오토인코더를 사용하였다. Poisson 손실 함수를 이용하여 100회의 epoch 동안 학습하였다(그림 1 참조).

오토인코더는 인코더 신경망과 디코더 신경망을 이용하여 데이터의 특징을 추출하는 대표적인 비지도 학습 인공 신경망 구조이다. 본 연구에서는 인코더 GCN 신경망을 통해 고차원 입력 데이터를 저차원 잠재 공간으로 변환(잠재 벡터 z)하고 이를 다시 디코더 신경망으로 복원한다. 생성된 잠재 벡터 z 는 고차원 입력 데이터의 주요 특징을 저차원으로 표현한다. 본 연구에서는 16 잠재 벡터를 사용하였다.

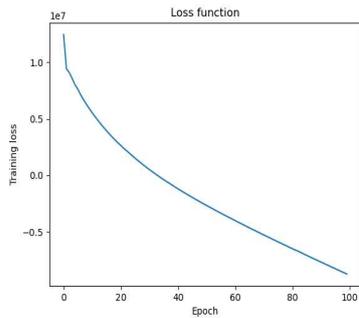


그림 1 손실 함수 그래프. y축은 단일 스칼라 값.

2.4 ASD 환자의 주요 뇌 연결성 도출

ASD 환자의 주요 뇌 연결성을 도출하기 위해 모든 뇌 연결성 12,720개를 하나씩 조정하여 학습된 GCN 기반 AE 모델에 입력, 출력된 결과값의 변화를 분석한 후, 가장 큰 변화를 유도하는 연결성(즉, 학습된 모델에 영향력이 큰 뇌 연결성)을 도출하였다. 각 연결성을 조정하기 위해 ASD 환자의 연결성 평균값에 표준편차의 3배를 더한 값을 적용하였다.

3. 결과

그림 2는 12,720개의 뇌 연결성에 대해 학습된 GCN-AE 모델에 미치는 영향력을 높은 순으로 정렬한 결과를 보여준다. 1-3번째 연결성과 4-5번째 연결성 사이에 간격이 있는 것을 확인할 수 있다. 영향력이 가장 높은 3개의 연결성은 기저핵(Basal Ganglia)과 배측 전두엽(Dorsal Frontal Cortex) 연결성, 방추회(Fusiform Gyrus)와 측소뇌(Lateral Cerebellum) 연결성, 그리고 전방 전두

피질(Anterior Prefrontal Cortex)과 하두정소엽(Inferior Parietal Lobule) 연결성이다.

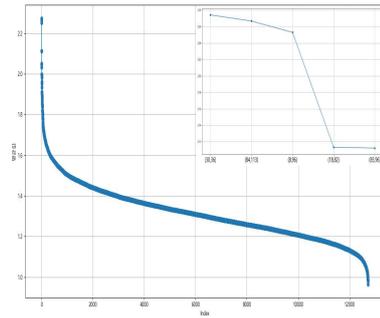


그림 2 학습된 GCN-AE 모델에 미치는 뇌 연결성

4. 논의

본 연구는 ASD 환자의 뇌영상을 GCN을 활용하여 비지도 학습하고, 학습된 모델 해석을 통해 ASD 환자의 주요 뇌 연결성을 도출하였다. 그 결과 ASD 환자의 주요 뇌 연결성으로 기저핵과 배측 전두엽 연결성, 방추회와 측소뇌 연결성, 그리고 전방 전두피질과 하두정소엽 연결성을 도출하였다.

ASD 환자와 같이 정신질환 환자를 겉으로 표현되는 증상을 기반으로 진단 또는 분류할 경우 진단 결과가 생물학적 특성과 일치하지 않을 수 있고, 주관적 견해가 반영될 가능성이 높다. 따라서 fMRI와 같이 생물학적 데이터에 기반한 진단 기법들이 연구되고 있다[2]. 본 연구에서는 GCN 모델을 활용하여 복잡한 뇌 네트워크 특성을 분석하고 오토인코더와 결합하여 데이터 기반 분석을 수행하였다.

학습된 모델 해석을 통해 뇌 영역 간의 주요 연결성을 추출하였으며, 모델에서 주요하게 다루어지는 입력 파라미터 즉, 3개의 뇌 연결성을 확인할 수 있었다. 해당 영역이 실제로 ASD와의 연관성이 높은지 또는 다른 정신질환 환자와 연관이 있는지 여부는 후속 연구를 통해 검증이 필요하다. 제안된 모델은 대규모 임상 데이터를 바탕으로 개발되었으며, 이는 기존의 증상 중심의 진단 방식을 넘어서 생물학적 데이터를 근거로 한 새로운 정신질환 분류 방법의 확립에 기여할 것으로 기대된다.

참고문헌

[1] Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The lancet*, 392(10146), 508-520.
 [2] Hyde, K. K., Novack, M. N., LaHaye, N., Parlett-Pelleriti, C., Anden, R., Dixon, D. R., & Linstead, E. (2019). Applications of supervised machine learning in autism spectrum disorder research: a review. *Review Journal of Autism and Developmental Disorders*, 6, 128-146.
 [3] Autism Brain Imaging Data Exchange (ABIDE), "Autism Brain Imaging Data Exchange (ABIDE)," 2024. [Online]. Available: https://fcon_1000.projects.nitrc.org/indi/abide/. [Accessed: Aug. 28, 2024].

수면 지표를 활용한 신체활동 예측: 스마트워치를 활용한 탐색적 연구1)

정현우¹, 손상진¹, 김예은², 이지원², 조한규^{1,2,*}

1. 국립군산대학교 컴퓨터소프트웨어특성화대학 인공지능융합학과

2. 국립군산대학교 전자정보공학부 정보통신전파공학전공

e-mail: jhw010608-@naver.com, thstkdwls13@naver.com, kimye0729@kunsan.ac.kr,

ljwjw18@kunsan.ac.kr, hgjo@kunsan.ac.kr*

Sleep Quality correlates with Physical Activity: An Exploratory Study Utilizing a Smartwatch

Hyunwoo Jung¹, Sangjin Son¹, Ye-eun Kim², Jiwon Lee², Han-Gue Jo^{1,2,*}

1. Department of AI Convergence, College of Computer and Software, Kunsan National University

2. School of Electronic and Information Engineering, Kunsan National University

요 약

최근 중대재해 처벌 등에 관한 법률 시행으로 인해 근로자 안전 관리의 중요성이 증가하면서, 본 연구는 스마트워치의 수면 데이터를 통해 근로자의 신체활동 지수를 예측하여 사고 위험 요소를 사전에 식별하고 예방할 수 있는 방안을 모색하였다. 연구는 스마트워치가 제공하는 수면 데이터와 개인이 자각하는 신체적 및 정신적 상태 간의 연관성을 규명하고, 이를 통해 다음 날의 신체적 활동을 예측할 수 있는지를 조사하기 위해 건강한 성인 1명을 대상으로 반복 측정을 실시하여 탐색적 연구를 수행하였다. 스마트워치가 제공하는 수면 지표와 개인이 자각하는 신체적 및 정신적 상태 간의 높은 연관성을 보였으며, 특히 수면점수 지표는 근력 운동 지표와 높은 상관성을 나타냈다. 이러한 결과는 스마트워치의 수면 데이터를 통해 근로자의 신체적 및 정신적 상태를 객관적으로 평가하고, 이를 기반으로 신체적 활동을 예측할 수 있는 가능성을 제시하였다. 향후 연구에서는 다양한 인구 집단과 장기적 데이터를 통해 연구 결과의 일반화 가능성을 높이고, 스마트워치 데이터를 활용한 안전사고 예방 시스템 개발을 개발할 필요가 있다.

1. 서 론

최근 중대재해 처벌 등에 관한 법률 시행으로 인해 근로자 안전에 대한 관심이 크게 높아졌다[1]. 중대재해처벌법은 경영책임자, 즉 사업주나 경영진이 산업재해에 대해 직접적인 책임을 지도록 규정하고 있으며, 경영진이 안전한 작업 환경을 조성하기 위해 더 큰 노력을 기울이도록 환경을 조성하고 있다. 또한, 이 법의 시행으로 인해 기업들은 기존의 안전 관리 시스템을 전면적으로 재검토하고 강화할 필요성이 커졌다. 안전 교육의 체계화, 위험 요소의 선제적 제거, 안전 장비의 확충 등 다양한 조치들이 적극적으로 이루어지고 있다. 특히, 근로자의 정신건강 관리와 같은 예방적 조치의 중요성도 부각되고 있다.

반복적이거나 고위험 작업을 수행하는 근로자의 경우, 수면 부족은 큰 위험 요소가 된다. 수면이 충분히 이루어지지 않으면 집중력, 판단력, 그리고 반응 속도가 저하되어 기계적 사고나 실수 발생 확률이 크게 증가한다[2]. 따라서 전날 수면의 질을 객관적으로 평가하고 이를 통해

근로자의 신체적, 정신적 상태를 확인하며, 사고를 미연에 방지하는 시스템 개발이 중요하다.

본 연구는 근로자의 신체적 및 정신적 상태를 사전에 예측하여 사고 위험 요소를 어떻게 선제적으로 제거할 수 있는지에 대한 통찰을 제공하는 것을 목표로 한다. 이를 위해 일상생활에서 생체신호를 측정하고 수면의 질을 평가할 수 있는 스마트워치를 활용하여, 수면 지표를 기반으로 다음 날의 신체활동(예: 근력 운동)을 예측할 수 있는지를 평가하였다. 특히, 스마트워치에서 제공하는 객관적 수면 지표와 개인이 자각하는 신체적 및 정서적 상태 간의 관계를 규명하고, 이러한 관계가 안전사고 예방 시스템 개발에 어떻게 활용될 수 있는지를 탐색한다.

2. 방법

2.1 데이터 수집

본 연구는 스마트워치가 제공하는 수면 데이터와 개인이 자각하는 신체적 및 정신적 상태 간의 연관성을 규명하고, 이를 통해 다음 날의 신체적 활동을 예측할 수 있는지를 조사하기 위해 건강한 성인 1명(나이: 23세, 성별: 남)을 대상으로 반복 측정을 실시하여 탐색적 연구를 수행하였다. 연구 기간 동안 삼성전자 갤럭시 워치 6를 이용하여 총 39일 간의 수면 지표를 측정하였으며, 5점 척도를

1) 본 연구는 2024년도 산업통상자원부 및 한국산업기술진흥원(P0012769, 2024년 산업혁신인재성장지원사업), 2024년도 과학기술정보통신부 및 정보통신기획평가원(SW중심대학사업 2023-0-00065)의 지원을 받아 수행되었음.

사용하여 수면 후 다음 날 아침의 신체 상태(피곤-상쾌), 기분 상태(불행-행복), 수면 상태(얕은 수면-깊은 수면)에 대한 주관적 자각 정도를 수집하였다. 여기서 1에 가까울수록 피곤, 불행, 얕은 수면을 나타내고, 5에 가까울수록 상쾌, 행복, 깊은 수면을 나타낸다. 연구 기간 중 39일 중 18일은 헬스장에서 근력 운동을 수행하였으며, 동일한 갤럭시 위치 6를 이용하여 심박수를 측정하였다.

갤럭시 위치를 통해 수집된 수면 관련 지표는 총 12가지로, 다음과 같다: 수면 시간(분), 수면 점수, 신체 회복 정도(%), 숙면 정도(%), 정신 회복 정도(%), 수면 주기(회), 수면 단계(awake)(%), 수면 단계(REM)(%), 수면 단계(light)(%), 수면 단계(deep)(%), 수면 중 평균 심박수(bpm). 운동 관련 지표는 총 9가지로, 다음과 같다: 운동 시간(분), 총 칼로리(kcal), 평균 심박수(bpm), 최대 심박수(bpm), 저항도 운동(93-112bpm)(%), 체중 조절 운동(113-130bpm)(%), 유산소 운동(131-149bpm)(%), 무산소 운동(150-149bpm)(%), 최대 심박수 운동(169-187bpm)(%).

2.2 데이터 분석

수면 지표, 운동 지표, 그리고 주관적으로 자각하는 신체적 및 정신적 상태(신체 상태, 기분 상태, 수면 상태) 간의 연관성을 규명하기 위해 피어슨 상관분석을 수행하였다. 통계 분석에는 SPSS 26 프로그램을 사용하였으며, 양측 검증을 통해 $p < 0.01$ 를 기준으로 통계적 유의성을 평가하였다.

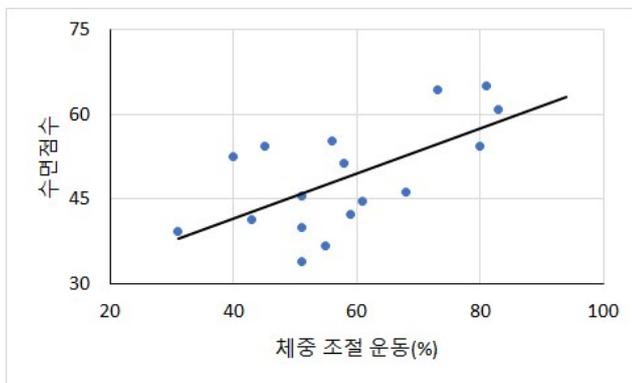


그림 1 수면 지표와 운동 지표의 상관관계

3. 결과

우선 갤럭시 위치를 통해 수집된 수면 관련 지표와 주관적으로 자각하는 신체적 및 정신적 상태와의 연관성을 분석하였다. 그 결과 주관적으로 자각하는 신체 상태, 기분 상태, 수면 상태는 각각 수면시간 ($r=0.642, p < 0.01$; $r=0.582, p < 0.01$; $r=0.686, p < 0.01$), 수면점수($r=0.725, p < 0.01$; $r=0.685, p < 0.01$; $r=0.817, p < 0.01$), 신체회복정도($r=0.706, p < 0.01$; $r=0.579, p < 0.01$; $r=0.790, p < 0.01$), 정신회복정도($r=0.615, p < 0.01$; $r=0.552, p < 0.01$; $r=0.600, p < 0.01$), 수면주기($r=0.553, p < 0.01$; $r=0.494, p < 0.01$;

$r=0.557, p < 0.01$), 수면단계_REM($r=0.498, p < 0.01$; $r=0.461, p < 0.01$; $r=0.461, p=0.01$)와 높은 양의 상관관계를 보였다. 이러한 결과는 스마트워치가 제공하는 수면 데이터와 개인이 자각하는 신체적 및 정신적 상태 간의 높은 연관성 나타낸다.

다음으로 수면 지표와 운동 지표 간의 연관성을 분석한 결과, 수면 점수와 체중 조절 운동 지표 간에는 높은 양의 상관관계가 나타났다($r=0.628, p < 0.007$; 그림 1 참조). 그러나 나머지 지표들 간의 연관성은 유의한 수준에 도달하지 못하였다($p > 0.01$).

4. 논의

본 연구는 객관적인 수면 지표와 주관적으로 자각하는 신체 상태, 기분 상태, 수면 상태 간의 관계를 분석함으로써, 스마트워치가 제공하는 객관적 수면 데이터와 개인의 주관적 자각 정도 간의 연관성을 규명하였으며, 이러한 데이터가 근로자의 안전관리 시스템에 어떻게 활용될 수 있는지에 대한 통찰을 제공한다.

수면 지표와 주관적으로 자각하는 신체적 및 정신적 상태 간에는 상당한 양의 상관관계가 나타났다. 특히, 수면 시간, 수면 점수, 신체 회복 정도, 정신 회복 정도, 수면 주기, 수면 단계(REM) 등은 주관적으로 자각하는 신체 상태, 기분 상태, 수면 상태와 강한 상관관계를 보였다. 이러한 결과는 스마트워치가 제공하는 객관적 수면 데이터가 개인의 주관적 자각과 밀접하게 연결되어 있음을 시사한다. 특히, 수면 점수와 신체 회복 정도, 정신 회복 정도는 개인의 주관적 평가와 높은 상관관계를 보였으며, 이는 수면의 질이 개인의 신체적 및 정신적 회복에 중요한 영향을 미친다는 것을 나타낸다.

수면 지표와 운동 지표 간의 연관성 분석 결과, 수면 점수와 체중 조절 운동 지표 간에는 높은 양의 상관관계가 발견되었다. 이는 수면점수 지표가 다음 날의 신체활동에 긍정적인 영향을 미친다는 것을 시사한다. 수면이 양호할수록 체중 신체활동에 대한 참여도가 높아질 수 있다는 결과는 근로자의 신체적 활동을 예측하는 데 활용할 수 있다. 즉, 전날 수면 중 스마트워치를 통해 측정된 수면 점수는 근로자의 전반적인 신체적 활동 수행 능력을 예측할 수 있을 것으로 기대된다.

참고문헌

[1] 최정학. (2021). 중대재해처벌법-기업 경영자 처벌의 논리. 노동법연구, (51), 1-38.
 [2] 이현정, & 김린. (2003). 만성수면부족의 사회적 영향. 수면정신생리, 10(2), 77-83.
 [3] Garbarino, S., Guglielmi, O., Sanna, A., Mancardi, G. L., & Magnavita, N. (2016). Risk of occupational accidents in workers with obstructive sleep apnea: systematic review and meta-analysis. Sleep, 39(6), 1211-1218.

LLaMA3 모델 기반의 Private LLM 시스템 설계

정도윤*, 이형호**, 정희자***, 김남호*

*호남대학교 컴퓨터공학과, ** (주)가민정보시스템, *** (주)휴넷가이아
e-mail : rabbit.white@daum.net, hhlee@comin.com, hngaia@naver.com,
nhkim@honam.ac.kr

Designing a Private LLM System Based on the LLaMA3 Model

DoYoon Jeong*, HyungHo Lee**, HeeJa Jeong***, NamHo Kim*

*Dept. of Computer Engineering, Honam Univ,

Comin Information Systems Inc, *HunetGaia Co.,Ltd

요약

본 논문은 LLaMA3(Large Language Model Meta AI)모델을 기반으로 한 Private LLM(Large Language Model) 시스템 구축 방법을 제안하며, LLM의 한계점인 맥락 부족과 환각 증상 등의 문제를 해결하기 위한 방안을 제시한다. ChatGPT 등 상용화된 모델을 사용하지 않고, 폐쇄망에서 생성형 모델을 작동시킬 수 있는 sLLM 모델을 구현하는 방법에 대해서 모색하고자 한다.

LLaMA3 모델은 페이스북에서 2024년에 공개한 오픈소스 LLM 모델 중 하나로 오픈 AI의 GPT-4o와 엔트로픽의 클로드 3(Claude 3)과 같은 모델이라고 볼 수 있다. 보안을 요구하는 비즈니스 환경의 영리 목적 회사에서 자체 모델을 설계하는데 활용할 수 있을 것이다. 결론적으로 본 연구는 LLaMA3 모델을 활용한 Private LLM 시스템의 설계와 구현에 대한 체계적인 접근을 통해, 데이터 프라이버시를 유지하면서도 고성능의 언어 모델을 효율적으로 활용할 수 있는 방안을 제시한다.

1. 서론

최근 몇 년간, 대형 언어 모델(LLM)은 자연어 처리(NLP) 분야에서 많은 변화를 가져왔다 [3]. 특히, OpenAI의 GPT 시리즈와 같은 상용 모델은 다양한 응용 프로그램에서 광범위하게 사용되고 있으며, 그러나 이러한 상용 모델을 사용할 경우, 데이터 프라이버시와 보안 문제가 중요한 우려 사항으로 떠오르고 있다. 개인 정보 보호가 필수적인 환경에서는 외부 서버에 민감한 데이터를 전송하는 것이 위험할 수 있으며, 이는 특정 산업에서의 LLM 활용을 제한하는 요소로 작용한다.

이러한 배경 속에서, LLaMA3 모델은 페이스북이 2024년에 공개한 오픈소스 LLM 모델로서, 상업적 목적뿐만 아니라 개인정보보호가 중요한 환경에서도 효과적으로 활용될 수 있는 가능성을 지니고 있다. 기업들은 폐쇄된 환경에서 생성형 모델을 구현하고, 데이터 프라이버시를 유지하면서도 고성능의 언어모델을 활용할 수 있는 기회를 가지게 된다 [1]. 본 논문에서는 LLaMA3 모델을 기반으로 한 프라이빗 LLM 시스템 구축 방안을 제안하고, 이를 통해 데이터 프라이버시와 보안을 강화하는 방법을 탐구한다. 마지막으로, 향후 연구 방향과 개선점을 제시하여, 개인 맞춤형 LLM 시스템의 발전에 기여하고자 한다.

2. 관련 연구

현재 OpenAI 외에도 GPT 기반의 다른 모델로는 Google

DeepMind의 PaLM2가 있다 [3]. 이 모델은 Google의 챗봇 바드(Bard)의 기반이 되는 모델이다. GPT-4가 대부분의 성능 벤치마크에서 선도하고 있지만, 다른 모델들도 특정 작업에서 유사한 성능을 나타내며 생성 트랜스포머 기반 언어모델의 발전에 기여하고 있다. PaLM2는 2023년 5월에 출시되었으며, 다국어 지원과 추론 능력 향상에 중점을 두고 더 효율적인 계산을 위해 훈련되었다. OpenAI와 Google DeepMind의 PaLM2 모델은 상용화 모델로서 대중에게 학습모델이 공개되어 있지 않다. Meta AI는 2023년 2월과 7월에 각각 LLaMA1와 LLaMA2 시리즈 모델을 출시했다. 이 모델들은 최대 70억 개의 매개변수를 갖추고 있으며, 오픈소스 모델로 공개함으로써 LLM의 기술 발전에 큰 영향을 미쳤다. LLaMA2는 다양한 매개변수 크기(7B, 13B, 34B, 70B)의 변형이 출시되었다. LLaMA는 비상업적 라이선스로 출시되었지만, LLaMA2는 연구와 상업적 사용을 위해 일반 대중에게 공개되었다. 2024년 4월, Meta AI는 기존 LLaMA2를 개선한 LLaMA3 버전을 출시하였으며, 이 모델의 학습 데이터셋 크기는 15T 토큰으로, LLaMA2에 비해 7배 이상의 향상된 데이터셋 크기를 제공한다 [1].

LLaMA2의 7B 및 13B 버전은 MHA(Multi-Head Attention) 구조를 가지고 있는 반면, 34B 및 70B 버전은 GQA(Grouped Query Attention) 구조를 채택하고 있다. LLaMA3는 GQA 구조를 기반으로 하여 추론 효율성을 개선하였다.

3. 제안 방법

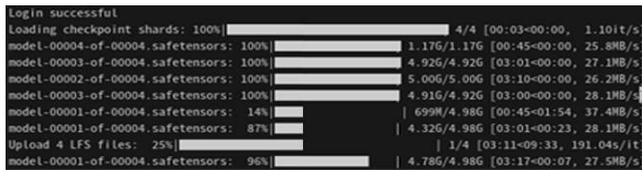
3.1 LLM 학습 모델 개발

제안된 방법은 기업 내부의 사전 지식 정보를 바탕으로 사전 데이터셋을 구축하는 것이다. 그러나 이러한 사전 구축 데이터셋의 정보는 오래된 지식이 될 가능성이 높다는 단점이 있다. 따라서 이러한 한계점을 이해하고, 애플리케이션 개발에 필요한 훈련 데이터를 생성하는 것이 중요하다.

	A	B	C	D	E	F	G	H
1	inputs		response					
2	rabbitsun2에서는 무엇을 가르치나요?		조보자 대상으로 빅데이터, 인공지능과 관련된 콘텐츠를 가르치고 있습니다.					
3	rabbitsun2는 누가 운영하나요?		대학원생 컴퓨터공학과 도움이 운영합니다.					
4	컴퓨터공학과는 무엇을 가르치나요?		컴퓨터공학과는 빅데이터, 인공지능, 웹개발 및 불특정인을 가르치고 있습니다					
5	컴퓨터공학과 등록비용은 얼마인가요?		등록비용은 국비지원 과정으로 무료 인니다.					
6	컴퓨터공학과는 1년에 몇 명을 선발하나요? 1년에 한번을 운영하고 있고 최대 100명을 선발합니다.							

(그림 1) 미리 정의된 데이터셋 구축

Hugging Face를 통해 Meta AI의 LLaMA3-instruct 모델을, 외부망을 통해 읽어 들인 후 고성능 GPU 시스템을 통해 학습한다 [1]. 이후에 생성된 데이터에 대해서 자연 어처리에 대한 fine-tuning 작업을 거친 후 로컬 모델에서 사용할 수 있는 형태로 변환 작업을 수행한다. 이때 시스템에서 생성되는 모델의 데이터량은 최소 10GB~100GB 이상의 대용량 파일이 생성된다. 데이터셋의 양에 따라서 훈련 데이터셋 크기는 달라질 수 있다.



(그림 2) 개발 환경에서 메타 모델에 필요한 최소 용량

3.2 애플리케이션

본 연구에서는 생성된 LLM 모델을 활용하여 로컬 시스템에서 OLLaMA를 통해 학습모델에 응답할 수 있는 구조를 설계하는 방법을 제안한다. 2022년 해리슨 체이스(Harrison Chase)에 의해 개발된 LangChain은 LLM 기반 애플리케이션 구축을 위한 오픈소스 파이썬 프레임워크이다 [2]. LangChain은 재사용할 수 있는 구성요소와 사전 구성된 체인을 제공하여 복잡한 LLM 응용 프로그램을 개발을 간소화한다. 체인(Chain)은 LangChain에서 모듈형 구성요소를 재사용할 수 있는 파이프라인으로 조합하는 개념을 의미한다. 가장 기본적인 체인은 PromptTemplate으로, 이는 형식화된 응답을 언어모델에 전달하는 역할을 한다.

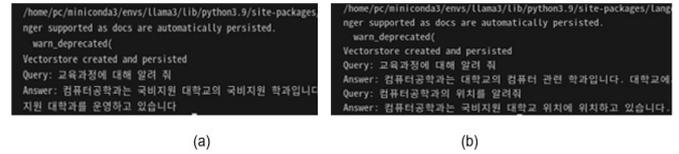
```

TEMPLATE """{%- if .System %}
<>{{ .System }}</>
{{- end %}}
<>Human:
{{ .Prompt }}</>
<>Assistant:
"""
SYSTEM """A chat between a curious user and an art
PARAMETER temperature 0
PARAMETER num_predict 3000
PARAMETER num_ctx 4096
PARAMETER stop <>
PARAMETER stop </>
    
```

(그림 3) 템플릿 구조 모델 형식

그림 3은 Template Structure Model 형식의 LLaMA3

모델을 사용하여 생성된 학습모델의 구조를 나타낸다. 그림 4는 LLaMA3 모델을 활용한 사용자 정의 학습 결과에서 얻은 결과를 나타낸다. 그림 4(a)는 환각의 예를 보여주고, 그림 4(b)는 정상 결과의 출력을 보여준다.



(그림 4) LLaMA 3 모델을 사용한 맞춤형 교육 결과

4. 결론

LLaMA3 모델을 사용하여 특정 요구사항을 충족하는 맞춤형 시스템을 개발할 수 있다. LLM의 한계점은 오래된 지식에 의존해서 최근 실세계의 정보를 제공하기 어렵다. 맥락 부족과 환각 위험이 있다. 이러한 문제점을 완화하는 방법으로 외부 컨텍스트를 제공하여 검색 증강을 한다. 프롬프트 공학과 체이닝, 모니터링과 필터링 리뷰, 파인 튜닝 등의 방법을 적용하여 증상을 완화할 수 있다 [2]. 현재 해외의 많은 연구자들은 LLaMA 모델을 바탕으로 활발한 연구를 진행하고 있으며 [4], 법률, 의료 등 특정 분야에서 활용할 수 있는 소규모 언어모델(Small Large Language Model, sLLM)의 활용성에 대한 다양한 방식으로 연구되고 있다. 그러나 국내에서는 소수의 대형 IT 기업을 중심으로 LLM 연구가 제한적으로 이루어지고 있어, 일반 연구자들에게는 연구 기회가 부족한 실정이다. 중소기업에서 활용할 수 있는 sLLM의 구축시 적용 가능할 것으로 기대한다.

Acknowledgement

본 논문은 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-002)

참고문헌

- [1] Dubey, A., Jauhri, A., Pandey, A., ... & Ganapathy, R, "The LLaMA 3 herd of models," arXiv preprint, arXiv:2407.21783, 2024.
- [2] Ben, Auffarth, 이병욱, LangChain으로 구현하는 LLM, 에어콘출판, p. 69-100, 2024.
- [3] Kucharavy, A, "From Deep Neural Language Models to LLMs," In Large Language Models in Cybersecurity: Threats, Exposure and Mitigation, pp. 3-17, 2024.
- [4] HoTaek, Joo., Sungha, Lee., KyungJoong, Kim., "인간과 ChatGPT의 대화내용을 이용한 공개 대형 언어모델 LLaMA 한국어 대화 능력 개선", 한국정보과학회 논문집, Vol. 2023 No. 6, 2023

YOLO를 통한 이미지 분석 기반 유실물 탐지 시스템 설계

조민성, 김민재, 김진성, 박정수, 김민서, 민홍
가천대학교 AI·소프트웨어학부

e-mail : cho010105@gachon.ac.kr, minjae20000514@gmail.com,
jskim6335@gachon.ac.kr, a62627629@gmail.com, mskim8055@gachon.ac.kr,
hmin@gachon.ac.kr

Image analysis-based lost and found detection system design using YOLO

Minseong Cho, Minjae Kim, Jungsu Park, Jinsung Kim, Minseo Kim, Hong Min
School of Computing, Gachon University

요약

본 연구에서는 공공장소에서 발생하는 유실물 탐지를 위한 MQTT 통신 기반 실시간 객체 탐지 및 위치 정보 전송 시스템을 제안한다. YOLOv5 모델을 사용하여 유실물 탐지 시스템을 개발하였으며, 실험에서는 서울시 공공자전거인 따릉이를 대상으로 하여 모델을 제작하였다. 라즈베리파이와 카메라 모듈을 활용하여 실시간으로 유실물(따릉이)을 탐지하였고, 탐지된 객체의 이미지를 서버로 전송하였다. 서버에서는 추가적으로 이미지 분석을 통해 탐지된 유실물의 위치와 고유 ID를 파악하고, 이를 데이터베이스에 저장하여 웹에서 시각화하였다. 실험 결과, 제안된 시스템은 높은 정확도와 실시간성을 보여주었으며, 유실물 관리에 효과적으로 적용될 수 있음을 입증하였다.

1. 서론

유실물의 발생은 공공장소에서 자주 발생하는 문제로, 이를 신속하게 찾아내고 관리하는 것은 시민의 편의와 공공시설의 운영 효율성을 높이는 데 중요한 역할을 한다. 공항, 기차역, 쇼핑몰과 같은 다양한 공공장소에서는 여러 종류의 유실물이 발생할 수 있으며, 이를 효과적으로 탐지하고 관리하기 위해서는 유실물의 기준점을 정의하고 회수할 수 있는 정밀한 기술이 필요하다. 기존의 유실물 관리 시스템은 주로 수동적인 방법에 의존하고 있으며, 유실물 발생 후 장시간이 지나야 탐지 및 회수가 이루어지는 경우가 대부분이다. 이는 유실물 회수의 성공률을 낮추고, 불필요한 비용과 시간이 소요되게 한다.

최근에는 컴퓨터 비전 기술의 발전과 함께 객체 탐지 알고리즘이 크게 발전하였으며, 그 중에서도 YOLO (You Only Look Once) 모델은 실시간 객체 탐지에 있어서 높은 정확도와 속도를 제공한다는 점에서 주목받고 있다. 특히, YOLOv5 모델은 이전 버전에 비해 경량화되고 성능이 향상되어 다양한 응용 분야에서 사용될 수 있다[1]. 본 연구는 YOLOv5를 활용하여 공공장소에서 발생하는 유실물을 실시간으로 탐지하고, 탐지된 유실물의 위치를 카메라로 파악한 뒤 MQTT 통신을 통해 서버에 전송하는 시스템을 제안하고자 한다.

본 연구에서는 실제 유실물 탐지를 위해 서울시 공공자전거인 따릉이를 유실물의 대체 객체로 사용하여 실험을 진행하였다. 이는 따릉이의 크기와 형태가 일정하며, 공공장소에서 흔히 발견될 수 있는 객체라는 점에서 모델의

성능을 평가하기에 적합하기 때문이다. 실험을 통해 수집된 따릉이 이미지 데이터는 라벨링 과정을 거쳐 YOLOv5 모델의 학습에 사용되었으며, 학습된 모델은 라즈베리파이 4와 카메라 모듈을 통해 실시간으로 따릉이를 탐지하는데 사용되었다. 탐지된 따릉이 이미지는 MQTT 프로토콜을 사용하여 서버로 전송되었고, 서버에서는 추가적인 분석을 통해 탐지된 따릉이의 위치와 고유 ID를 파악하여 Firebase에 그 데이터를 저장하였다. 이 정보는 웹에서 시각화되어 실시간으로 회수되지 않은 따릉이의 위치를 모니터링할 수 있도록 하였다.

2. 설계

2.1 모델 학습

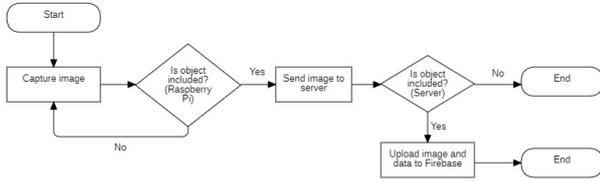
실험에서는 서울시 공공자전거 따릉이를 유실물의 대체 객체로 사용하였다. 따릉이와 비따릉이(다른 자전거)의 이미지를 약 200장 수집하여, 1:1 비율로 라벨링을 진행하였다. 수집된 데이터는 8:2 비율로 트레인 셋과 검증 셋으로 나누어 YOLOv5 모델을 학습시켰다.

YOLOv5 모델은 Pytorch 기반으로 학습되었으며, 학습된 모델은 TensorFlow Lite를 사용하여 변환되어 라즈베리파이에서 실시간으로 사용할 수 있도록 하였다. 학습 과정은 50epoch 동안 진행되었으며, 검증 셋에서 약 95%의 정확도를 달성하였다.

2.2 객체 탐지 및 서버 전송

라즈베리파이는 실시간으로 유실물(따릉이)을 탐지하는

역할을 수행한다. 이를 위해 라즈베리파이에는 카메라 모듈 3이 연결되어 있으며, 이 카메라는 영상을 캡처하여 YOLOv5 모델에 입력한다. 탐지된 유실물의 이미지는 MQTT프로토콜을 통해 서버로 되었다.



(그림 1) 객체 탐지 및 서버로의 전송 과정

2.3 서버(Server)

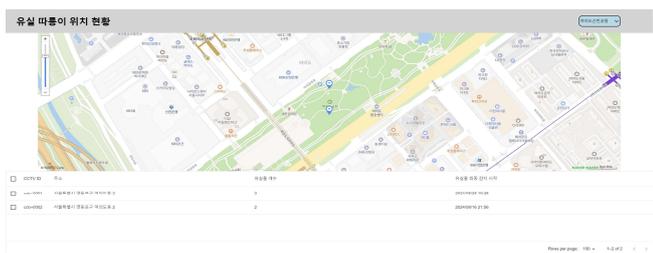
서버는 라즈베리파이로부터 전송된 이미지를 수신하고, 이를 다시 한번 분석하여 보다 정밀한 객체 탐지와 위치 추적을 수행한다. 서버에서는 YOLOv5 모델이 다시 구현되어 있으며, 수신된 이미지를 바탕으로 바운딩 박스를 생성한다. 처리된 객체의 이미지, 위치 정보, 고유 ID, 시간 정보를 추출하여 이를 데이터베이스에 저장할 수 있는 형식으로 변환한다. 서버는 또한 파이어베이스로 데이터를 전송하는 중간 역할을 하며, 실시간 데이터를 빠르고 안정적으로 처리할 수 있도록 최적화되었다.

2.4 파이어베이스(Firebase)

파이어베이스는 서버에서 출력하는 최종 산출물을 저장한 후, 권한을 가진 사용자에게 한해 데이터 제어를 허용하는 클라우드 데이터 저장소의 역할을 한다. 본 시스템에서는, 데이터를 저장할 작업을 파이어베이스에 이관하고, 서버는 그 외의 모델 학습 및 시스템 로직을 수행하도록 구성되었다.

2.5 웹 인터페이스(Web Interface)

앞선 데이터 처리 과정의 최종 결과물을 사용자에게 나타내는 웹페이지는, CCTV가 마커로 표시된 Naver Map API를 활용하여 서울시 내에 존재하는 공원 속 CCTV들이 서버로 전송한 유실 따릉이 데이터를 React 컴포넌트 중 하나인 MUI X Datagrid의 형태로 출력하였다. Datagrid 내에는 CCTV ID, 주소, 해당 CCTV가 감지한 유실물의 개수, 그리고 파악된 유실물의 최종 감지 시간을 출력하며, 해당 CCTV의 행을 클릭하여 해당 CCTV가 감지한 유실물의 최초 및 최종 발견 시각과 사진을 파악할 수 있다.



(그림 2) 유실 따릉이 위치 현황 안내 웹페이지

3. 결론

YOLOv5는 이미지 내에 존재하는 객체를 분석하는 데에 있어 높은 정확도와 빠른 속도를 제공한다. 본 논문은 YOLOv5 기반의 이미지 객체 탐지와 MQTT 통신을 활용하여 서울시 따릉이라는 유실물을 탐지하는 과정과, 해당 기술이 탑재된 CCTV가 탐지한 유실 따릉이를 사용자에게 나타내는 방식을 백엔드와 프론트엔드로 나누어 서술하였다.

참고 문헌

[1] J. R. Terven and D. M. Cordova-Esparza, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," arXiv preprint, 2023.

자폐스펙트럼 장애 진단 그래프 신경망 모델 비교 연구¹⁾

김다연^{1,†}, 백빈나^{1,†}, 배하람^{1,†}, 유채린^{2,†}, 이지원³, 조한규^{1,3,*}

1) 국립군산대학교 컴퓨터소프트웨어특성화대학 인공지능융합학과

2) 국립군산대학교 컴퓨터소프트웨어특성화대학 IT융합통신학과

3) 국립군산대학교 전자정보공학부 정보통신전파공학전공

e-mail : rlaekdus75@naver.com, bibphan@naver.com, bae664479@naver.com,

younanhiw@naver.com, ljwjl18@kunsan.ac.kr, hgio@kunsan.ac.kr*

Graph neural network models for autism spectrum disorder identification

Da-Yeon Kim^{1,†}, Bin-Na Beak^{1,†}, Ha-Ram Beak^{1,†}, Chae-rin Yu^{2,†}, Ji-Won Lee³, Han-Gue Jo^{1,*}

1) Department of AI Convergence, College of Computer and Software, Kunsan National Univeristy

2) Department of IT Convergence Communication Engineering, Kunsan National University

3) School of Electronic and Information Engineering, Kunsan National Univeristy

요 약

본 연구는 뇌영상을 기반으로 자폐 스펙트럼 장애(Autism Spectrum Disorders, ASD)를 구별할 수 있는 딥러닝 모델을 개발한다. 다양한 그래프 신경망(Graph Neural Network, GNN) 모델을 활용하여 뇌 영역 간의 복잡한 네트워크 활동 패턴을 분석하고 ASD 환자와 정상 대조군을 분류하였다. 연구에서 사용된 데이터셋은 Autism Brain Imaging Data Exchange(ABIDE)에서 제공된 883명의 기능적 자기공명영상(fMRI) 데이터를 기반으로 하였으며, Graph Convolutional Network (GCN), Graph Attention Network(GAT), Self-Attention Graph Pooling(SAGPool) 모델을 학습하였다. 각 모델은 ASD와 정상인을 분류하는데 약 74.1%, 78.9%, 92.6%의 정확도를 보였다.

1. 서 론

자폐 스펙트럼 장애(Autism Spectrum Disorder, ASD)는 생후 3년 이내에 발현되는 신경 발달 장애로 대뇌 피질의 신경회로 발달이 저하되거나 비정상적인 성장 패턴을 보이는 것이 주요한 병리적 특징으로 알려져 있다[1]. 따라서 뇌 영역 간 연결성 손상은 ASD의 주요 발병 원인 중 하나로 지목되고 있으며, 행동적·사회적 상호작용의 결핍과 제한된 반복적 행동이 주요 증상으로 나타난다.

최근 연구에서는 기능적 자기공명영상(fMRI) 데이터를 활용하여 병리적 특성 즉 뇌의 활동성을 분석하여 ASD와 정상인을 구분하는 기법들이 소개되고 있다[2]. 그러나 대규모 뇌 데이터에 기반하여 복잡한 뇌 네트워크 특성을 분석할 수 있는 새로운 접근법의 필요성이 제기되고 있다.

이에 본 연구는 대규모 fMRI 데이터와 복잡한 뇌 네트워크 특성을 분석할 수 있는 딥러닝 기술을 활용하여 ASD 환자와 정상인을 구분할 수 있는 딥러닝 모델을 개발하였다. 특히 뇌 영역 간 복잡한 연결성을 효과적으로 모델링하기 위해 그래프 신경망(Graph Neural Network, GNN)을 적용하였다. 뇌의 기능적 연결성을 그래프로 표

현하고 이를 GNN으로 분석함으로써, 기존의 기계학습 및 통계적 방법보다 향상된 분류 능력을 제공할 수 있을 것으로 기대한다.

2. 연 구

2.1 데이터 수집

본 연구에서는 Autism Brain Imaging Data Exchange (ABIDE)[3] 데이터를 통해 정상 대조군(Healthy Controls, HC) 475명과 ASD 환자 408명, 총 883명의 임상 데이터를 사용하였다. 휴지기 상태에서 Blood Oxygen Level Dependent(BOLD) 신호를 기반으로 뇌의 활성화 정도를 간접적으로 반영하는 fMRI 데이터를 사용하였다.

2.2 데이터 전처리

Dosenbach's 160개 뇌 영역[4]으로 세분화된 fMRI 데이터를 각 영역 간의 BOLD 시계열 데이터를 기반으로 상관 계수 행렬을 산출하였다. 산출된 상관 행렬(160x160)은 뇌 영역 간의 기능적 연결성을 수치화하여 모델의 입력 데이터로 사용되었다. 데이터의 일관성과 학습 과정의 안정성을 높이기 위해 상관값을 Fisher z-transformation하고 데이터의 결측값을 제거하였다. 최종적으로 394명의 ASD 환자와 461명의 정상 대조군 데이터를 포함한 학습용 데이터셋을 구축하였다.

1) 본 연구는 국립대학육성사업비, 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업”지원을 받아 수행되었음(2023-0-00065)

† These authors contributed equally to this work

2.3 실험방법

본 연구에서는 특정 뇌 영역에서 발생하는 복잡한 네트워크 활동 패턴을 분석하기 위해 Graph Convolutional Network (GCN), Graph Attention Network(GAT) 및 Self-Attention Graph Pooling(SAGPool)과 같은 다양한 GNN 모델을 활용하였다.

GCN 모델은 각 노드가 이웃 노드의 특징을 집계하여 자신의 특징을 갱신하는 방식으로 작동한다. 이 과정은 그래프 구조에 맞춘 컨볼루션 연산을 통해 이루어지며 그래프 데이터의 구조적 특성을 효과적으로 학습한다[5].

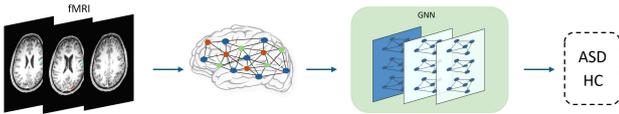


그림 1 데이터 분석 흐름도

GAT 모델은 그래프 데이터를 처리하기 위해 Self-Attention을 활용하는 신경망이다. Self-Attention을 활용해 이웃 노드 간의 중요도를 연산하고 각 이웃 노드에 대한 서로 다른 가중치를 할당한다[6].

SAGPool 모델은 GNN에서 각 노드의 중요도를 학습하기 위한 Self-Attention 기법, 그래프 크기를 축소하기 위한 Graph Pooling 기법, 축소 후 각 레이어의 특징을 결합하기 위한 JumpingKnowledge 기법을 활용한다[7]. 이를 통해 큰 그래프 구조 데이터를 효율적으로 처리하고, 중요도가 높은 노드만을 남겨 그래프 구조 압축 및 복잡성을 줄였다.

2.4 훈련과정

모델의 과적합을 방지하기 위해 dropout=0.3을 적용하였으며 학습데이터(69%), 테스트 데이터(20%), 검증데이터(11%)로 나누어 학습하였다.

GCN 모델은 Chebyshev Convolutional Layer (ChebConv)를 활용하여 그래프 데이터의 특징을 추출한다. 이 레이어는 K차 Chebyshev 다항식을 통해 각 노드의 특징을 이웃 노드들의 정보로 갱신하며, K_order은 그래프에서 고려하는 이웃의 범위를 결정한다. 본 연구에서는 K=3으로 설정하였다.

GAT 모델은 헤드의 수와 레이어 층 수, 은닉층 차원 수에 대해 다양한 경우의 수를 조합하여 가장 높은 정확도를 보이는 조합을 추출하였다.

SAGPool 모델은 레이어 층 수, 풀링 비율, 은닉층 차원 수를 조정하였다. 다른 모델에는 없는 ratio 파라미터는 pooling 단계에서 선택되는 노드의 비율을 결정하며 모델의 성능과 그래프 표현의 간결성에 영향을 미친다.

3. 결과 및 결론

모델별로 은닉층(hidden layer) 수에 따른 입력력 노드

의 개수를 조절하여 실험한 결과 GCN 모델의 평균 정확도는 $74.09 \pm 3.07\%$, GAT 모델은 $78.88 \pm 1.62\%$, SAGPool 모델은 $92.58 \pm 1.76\%$ 로 나타났다. 이 중 SAGPool 모델이 가장 높은 정확도를 보여주었다. GCN과 달리 GAT와 SAGPool은 Attention 기법을 사용하여 더 높은 정확도를 나타내지만, Multi-head Attention 방법을 사용하는 GAT는 그래프의 모든 정보를 병렬화하는 과정에서 중복계산이 발생할 수 있다. 그러나 SAGPool은 Self-Attention 및 Graph Pooling 기법을 활용하여 그래프를 간소화하면서 과적합을 방지한다. 이러한 SAGPool 모델 특성이 뇌의 복잡한 네트워크 구조를 보다 효율적으로 분석하여 ASD 환자와 정상인 대조군을 GCN이나 GAT 모델보다 정확하게 분류하는 것으로 해석할 수 있다.

(표 1) 모델별 정확도

Model	Accuracy	평균값	표준편차
GCN		74.09%	$\pm 3.07\%$
GAT		78.88%	$\pm 1.62\%$
SAGPool		92.58%	$\pm 1.76\%$

참고문헌

- [1] Lord, C., Cook, E. H., Leventhal, B. L., & Amaral, D. G. (2000). Autism spectrum disorders. *Neuron*, 28(2), 355-363.
- [2] Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17, 16-23.
- [3] Autism Brain Imaging Data Exchange (ABIDE), "Autism Brain Imaging Data Exchange (ABIDE)," 2024. [Online]. Available: https://fcon_1000.projects.nitrc.org/indi/abide/. [Accessed: Aug. 28, 2024].
- [4] Nico U. F. Dosenbach et al., Prediction of Individual Brain Maturity Using fMRI. *Science* 329,1358-1361(2010). DOI:10.1126/science.1194144.
- [5] Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1), 1-23.
- [6] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *arXiv preprint arXiv:1710.10903*.
- [7] Lee, J., Lee, I., & Kang, J. (2019, May). Self-attention graph pooling. In *International conference on machine learning* (pp. 3734-3743). pmlr.

디지털 트윈 도시에서 3D 객체와 실시간 스트림처리의 의미론적 통합

이예원, 이성범, 한정규, 천세진
 동아대학교 컴퓨터AI공학부

{2243986, sungbeom}@donga.ac.kr, {jkhan, sjchun}@dau.ac.kr

Semantic Integration of 3D Objects and Real-Time Streams in Digital Twin City

Yewon Lee, Seongbeom Lee, Jungkyu Han, Sejin Chun
 Department of Computer Engineering and AI, Dong-A University

요약

디지털 트윈 도시에서 중요한 도전 중의 하나는 다양한 이종의 가상 데이터를 실시간으로 통합하고 연결하는 것이다. 디지털 트윈 도시 분야의 기존 정보 모델은 도시 개체와 연관된 실시간 센싱 데이터, 복잡한 이벤트, 그리고 정보 품질을 표현정보가 결여되었다. 이러한 문제를 해결하기 위해, 본 논문에서는 가상의 3차원 디지털 트윈 도시 객체를 표현하고 관련된 개체 및 생성되는 실시간 스트림 데이터를 표현하는 의미적 정보모델 초안을 제시한다.

1. 서론

시맨틱 모델(Semantic model)은 다양한 산업의 도메인 지식을 통합하고 의미적인 상호운용(Semantic Interoperability) 및 일관성을 향상하기 위한 정보 모델로 주목받고 있다. 이러한 모델들은 산업 4.0 내 제조공정[1], 지속가능한 미래를 위한 빌딩[2] 내 에너지 절감, 디지털 트윈 기반 관리행정 등 다양한 영역에서 서비스나 이벤트, 실제 개체, 실시간 데이터, 정보 품질과 같은 다양한 정보들을 제공한다.

온톨로지(Ontology)는 시맨틱 모델을 표현하는 핵심 도구이고 현재 디지털 트윈 도시 분야에서 효과적인 상호운용성을 확보하기 위해 널리 사용되고 있다[3]. 온톨로지는 개념(Class), 속성(Property), 관계(Relationship), 인스턴스(Instance)로 구성 되어있다. 이러한 구성들로 그래프 구조를 만들어 도메인 내 개념들의 집합이나 그들 간의 관계를 구조화, 형식적으로 표현할 수 있다. 이 중, Web Ontology Language (OWL)는 대표적인 언어이다.

디지털 트윈 도시 분야에 적용하는 온톨로지는 다양한 산업의 정보들을 의미적으로 상호연결되어야 하고 정보들 간의 일관성을 유지해야한다. 하지만 기존 온톨로지는 이러한 관계가 정교하게 표현되지 못할뿐만 아니라 개체에 대한 정보는 스트림처리가 불가능하다. 따라서 본 논문은 이러한 문제점을 해결하기 위해 기존의 온톨로지를 재사용하여 도시 내의 다양한 개체들로부터 스트림데이터를 다루고 정교한 관계를 표현할 수 있는 온톨로지 초안을 제시한다.

2. 관련 연구

디지털 트윈 도시 분야 온톨로지에 대한 기존 연구는 도시 내 개체와 그들 간의 관계를 표현하는 데 관심이 있었다. ifcOWL온톨로지[4]는 Building Information Model(BIM) 프로세스를 위한 디지털 표준 온톨로지로서 데이터 모델 교환 명세, 커뮤니케이션 프로토콜, 기기정보 등을 포함한다. Brick온톨로지는[2] 빌딩 내 물리적, 논리적, 가상적 차원과 그들 간의 관계에 대해 표현한다. CityGML¹⁾ 모델은 3D도시 개체에 대한 표준적인 정보 모델로서 LOD(Level

of Detail) 5단계까지 도시 개체들에 대한 상세한 정보를 나타낸다. 특히, Appearance 계층에 Thematic 모듈은 Relief, Water Body, Vegetation, Land Use 등 feature에 대한 상세한 정보표현을 지원할 뿐만 아니라, Noise Emission과 같은 확장 모듈도 지원한다. EU의 SAREF온톨로지는[5] 스마트시티의 관심 대상(Feature of Interest), 서비스 기기, 운영에 대한 특징과 에너지/환경/빌딩 도메인에 대한 확장을 지원한다.

앞서 언급한 기존 연구들은 개체의 정보는 표현하지만, 디지털트윈 내 단일 또는 복잡한 개체들의 센싱 데이터로부터 파생되는 서비스 품질에 대한 정교한 관계 설정이 부재하다. 그리고 디지털 트윈 도시 개체에서 발생하는 스트림데이터에 대한 생성, 처리, 그리고 분석에 대한 표현이 결여되어 있다.

3. 제안 온톨로지

제안 온톨로지는 가상의 3차원 공간정보를 표현하고 개체들로부터 생성되는 데이터와 스트림데이터에 대한 개념과 관계를 제공한다. 또한, 서비스 및 이벤트에 관한 Complex Event Service(CES), 스마트시티 내 3D 개체인 CityGML, 데이터 스트림에 관한 SAO(Stream Annotation Ontology), 센서 관측 정보에 관한 Semantic Sensor Network(SSN), 데이터 품질에 관한 Quality-Of-Information(QOI)의 하위 온톨로지들로 재사용되었다. 개체 간 스트림데이터를 효과적으로 사용을 위해 하위 온톨로지의 주요한 개념과 관계를 재사용하고 그들 간의 상호연관된 관계를 표현한다.

3.1 제안된 모델에서 사용된 온톨로지와 기능

그림 1은 CityGML, SSN, CES, SAO 그리고 QOI 온톨로지를 재사용 및 통합하여 구성된 모델이다. 재사용된 온톨로지의 세부설명은 다음과 같다.

1) CityGML: <https://www.ogc.org/standard/citygml/>

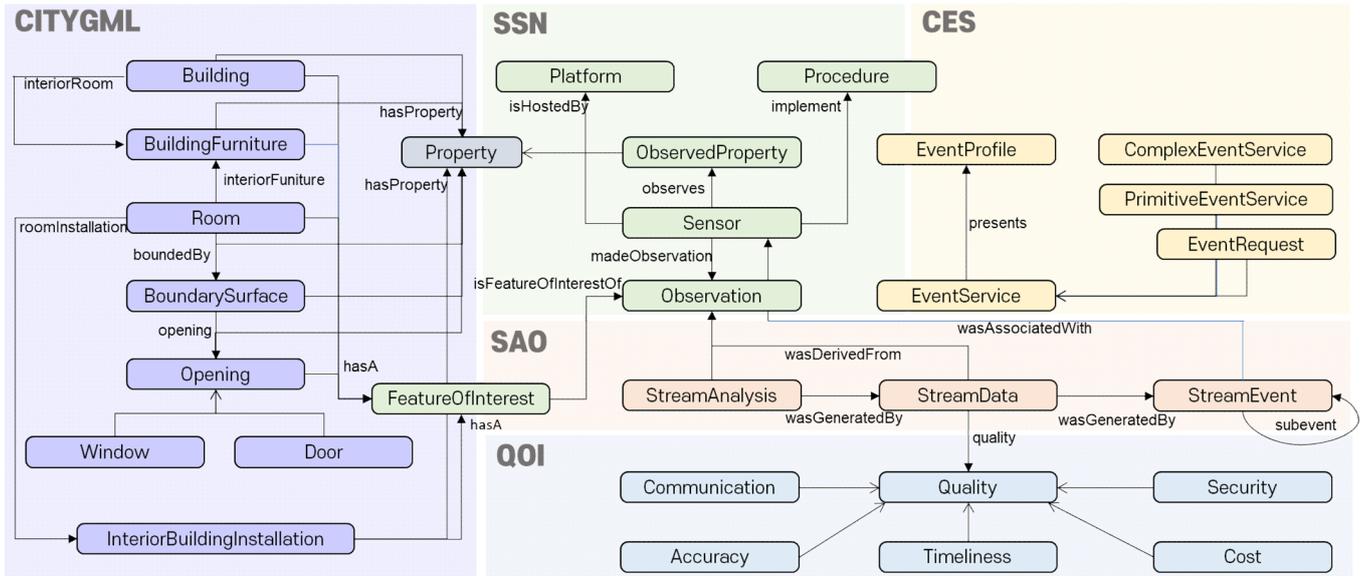


그림 1. 디지털 트윈을 위한 온톨로지 초고(Ontology draft)

- 1) CityGML: 빌딩(Building)의 빌딩 가구(BuildingFurniture)는 내부 요소(Room)에 대한 제한된 범위(BoundarySurface)의 정보(Opening)를 활용하고, 요구사항에 따라 확장가능하며, 공간적 관계 정보가 LoD4수준까지 표현이 가능하다.
- 2) SSN: 센서(Sensor)와 그들의 관측 정보(ObservedProperty), 센서가 탑재된 플랫폼(Platform), 사용되는 프로시저(Procedure)를 표현하여 보편적으로 사용되는 정보만을 표현하였으며, 다수의 관측을 통한(ssn:Phenomenon)에 대한 정의를 위한 SSN의 확장 및 호환이 가능하다.
- 3) CES: Event Ontology에 기반하여, 복잡한 이벤트서비스와, 단위 이벤트를 포함하는 이벤트 서비스(ces:EventService)를 정의한다. 해당 이벤트 서비스는 seo:StreamData에 의해 관측된다.
- 4) SAO: 시간적인 개념인 sao:StreamData, 관측의 스트림인 sao:StreamEvent, 데이터 분석과정인 sao:StreamAnalysis를 포함한다.
- 5) QoI: 디지털 트윈 내 데이터 스트림의 정보의 품질을 다루기 위해 사용되며 보안, 정확성 같은 핵심 품질은 스트림데이터(StreamData)의 필수적인 평가 되도록 돕는다.

3.2 온톨로지 간 공유되는 개념

다른 온톨로지를 하나의 모델로 통합하기 위해, 우리는 온톨로지 간 의미적으로 일치하는 개념을 매칭하는 온톨로지 정렬(Ontology alignment)을 수행하였다. 따라서 우리는 CityGML과 SSN를 연결하는 Property개념과 SSN과 SAO를 연결하는 Observation개념, 마지막으로 SAO와 QOI를 연결하는 StreamData개념을 사용한다. 각 개념들이 연결되어 하는 기능은 다음과 같다.

- 1) 속성(Property)은 CityGML에서 개체가 가지는 특징뿐만 아니라 SSN의 특정 센서에 의해 관측된 속성과 연관되어 있다. 관측된 속성은 연관된 관측 대상을 찾는다.
- 2) 관측(sao:Observation)은 SAO의 관심대상의 속성값을 추정하거나 계산한다.
- 3) 스트림데이터(StreamData)은 스트림분석(StreamAnalysis)을 생성한다. QOI의 데이터 품질을 다루고 있다.

3. 결론

우리는 디지털 트윈 도시를 나타내는 온톨로지 모델의 스트림 데이터 처리 문제와 개체의 센싱 데이터로부터 정교한 관계가 결여되어있는 문제에 대해 해결하고자 상호운용성이 높은 온톨로지 모델을 제안하였다. 제안한 온톨로지 모델은 이중 및 다양한 플랫폼 내 서비스와 정보의 품질을 표현하고 스트림데이터를 분석하고 처리할 수 있는 경량화 된 시맨틱 모델로 구체적인 디지털 트윈 도시 객체를 나타내는 데 사용될 뿐만 아니라 디지털 트윈과 같은 가상 세계를 나타내는 데 사용할 수 있다.

감사의 글

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가의 지원을 받아 수행된 연구임 (No. 2023-0-00076, SW중심대학(동아대학교))

참고 문헌

- [1] Buchgeher, G., Gabauer, D., Martinez-Gil, J., & Ehrlinger, L. (2021). Knowledge graphs in manufacturing and production: a systematic literature review. *IEEE Access*, 9, 55537-55554.
- [2] Fierro, Gabe, et al. "Formalizing tag-based metadata with the brick ontology." *Frontiers in Built Environment* 6 (2020): 558034.
- [3] Antoniou, Grigoris, and Frank van Harmelen. "Web ontology language: Owl." *Handbook on ontologies* (2009): 91-110.
- [4] Beetz, J., Van Leeuwen, J., & De Vries, B. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Ai Edam*, 23(1), 89-101.
- [5] Daniele, L., den Hartog, F., & Roes, J. "Created in close interaction with the industry: the smart appliances reference (SA REF) ontology", FOMI 2015, Proceedings 7 (pp. 100-112).

정보 취약 계층을 위한 블록체인 기반 노인 구직 시스템 설계

조민성¹, 김민서¹, 김민재¹, 이영찬¹, 민홍¹
¹가천대학교 AI·소프트웨어학부

cho010105@gachon.ac.kr, mskim8055@gachon.ac.kr, minjae20000514@gmail.com,
lych0918@gachon.ac.kr, hmin@gachon.ac.kr

Designing a blockchain-based senior job search system for the information-vulnerable

Minseong Cho¹, Minseo Kim¹, Minjae Kim¹, Youngchan Lee¹, Hong Min¹
¹School of Computing, Gachon University

요 약

블록체인은 네트워크 내에서 정보를 암호화하는 데 있어 매우 유용한 기술이다. 이러한 블록체인의 높은 보안성으로 인해 해당 기술은 암호화폐나 공공서비스 등 다양한 분야에서 연구되고 활용된다. 본 논문에서는 블록체인을 활용하여 온라인 정보 수집에 취약한 노인들을 대상으로 서비스하는 일자리 애플리케이션 시스템에 대한 설계를 제안하였다.

1. 서 론

고령화 사회로의 진입은 여러 사회적 과제를 수반하며, 특히 노인들이 적합한 일자리를 찾는 과정에서 디지털 정보 접근성 부족으로 인한 어려움이 증가하고 있다. 노인들의 경우 디지털 환경에 익숙하지 않아 온라인을 통한 일자리 정보를 활용하는 데 어려움을 겪으며, 이러한 문제는 점차 많아지는 고령자의 경제적 어려움과 사회적 고립을 심화시킬 수 있다. 더 나아가, 법정 대리인의 도움 없이 온라인 서비스를 이용하기 어려운 경우가 많아 공공기관은 고령자와 법정 대리인의 관계를 안전하게 관리할 수 있는 솔루션이 요구된다.

블록체인 기술은 이러한 문제를 해결할 수 있는 효과적인 기술로 현재 논의되고 있다. 블록체인은 데이터를 암호화하고 분산된 네트워크에 저장함으로써 데이터의 안전성과 무결성을 보장한다. 공공기관이 고령자와 법정 대리인의 정보를 안전하게 관리할 수 있는 이유는 블록체인의 암호화 및 분산 원장 기술 덕분이다. 데이터 전송 시, 비대칭 암호화를 통해 송신자의 공개키로 암호화된 데이터는 수신자만이 자신의 개인키로 복호화할 수 있으며, 이를 통해 중간에서 데이터가 탈취되거나 위조되는 것을 방지한다[1]. 또한, 블록체인은 한 번 기록된 데이터가 쉽게 수정되거나 삭제될 수 없는 변경 불가능성(Immutability)을 제공하여, 공공기관이 법정 대리인과 고령자의 인증 정보를 관리할 때, 블록체인에 기록된 정보는 위변조의 위험 없이 안전하게 저장되며, 네트워크의 모든 노드에 동일한 데이터가 기록된다. 이를 통해 중앙 집중식 시스템에서 발생할 수 있는 데이터 손실이나 보안 문제를 방지할 수 있다.

본 논문에서는 이러한 블록체인의 안전성을 바탕으로, 고령자 맞춤형 일자리 매칭 서비스를 제안한다. 이 시스템

은 고령자와 법정 대리인의 관계를 안전하게 관리하고, AI 기반의 맞춤형 일자리 추천 기능을 통해 고령자들이 적합한 일자리를 쉽게 찾을 수 있도록 돕는다. 블록체인을 활용한 데이터 보호와 AI의 효율적인 일자리 매칭이 결합된 본 시스템은 고령자들의 사회적 참여를 증진시키고, 공공기관이 고령자 서비스를 운영하는 데 있어 신뢰성과 투명성을 보장할 수 있을 것이다.

2. 설계

2.1 법정 대리인 인증 및 관리

고령자가 시스템을 이용하기 위해 법정 대리인이 필요한 경우, 시스템은 블록체인을 통해 대리인의 신원을 검증하고 관계를 인증하는 프로세스를 구현한다. 먼저 고령자가 대리인을 등록할 때, 해당 대리인의 정보는 암호화되어 블록체인에 기록된다. 등록된 정보는 대리인이 법정적으로 인증된 인물임을 증명하는 데이터와 연결되며, 블록체인에 분산 저장되므로 중앙 서버에서 관리되지 않는다. 이를 통해 공공기관은 대리인의 신원을 안전하게 관리할 수 있다. 또한 대리인의 인증을 위해, 고령자가 대리인의 도움을 받아 서비스를 이용할 때, 블록체인에 기록된 대리인 정보를 기반으로 대리인의 신원이 즉시 검증된다.

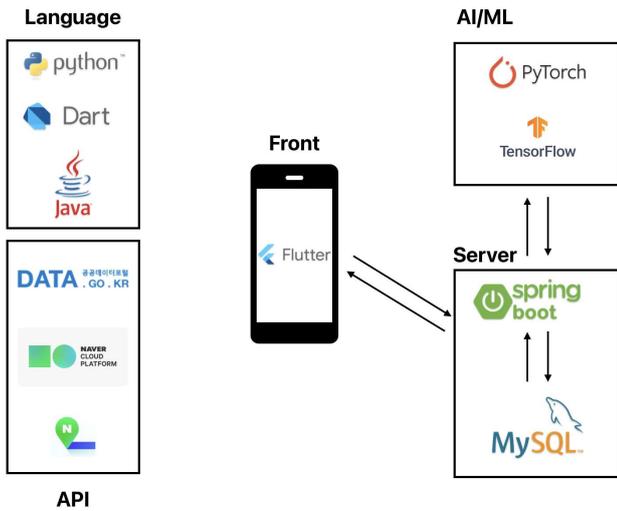
2.2 블록체인을 통한 일자리 매칭 데이터 보안

블록체인은 고령자의 개인정보 및 일자리 매칭 데이터를 안전하게 보호하는 데 핵심적인 역할을 한다. 먼저 고령자가 제공한 개인정보는 블록체인에 기록되며, 암호화된 상태로 안전하게 저장된다. 블록체인의 분산 원장 시스템을 통해 이러한 정보는 네트워크 내 여러 노드에 분산 저장되므로, 중앙 서버에서 발생할 수 있는 데이터 손실 및 해킹에 대한 위험을 줄일 수 있다. 이후의 일자리 매칭 결

과 및 고령자의 매칭 이력은 투명하게 블록체인에 기록된다. 고령자가 어떤 일자리에 지원했는지, 어떤 추천이 이루어졌는지 등 모든 정보가 시간 순서대로 기록되며, 고령자와 대리인이 원하는 경우 언제든지 과거 기록을 확인할 수 있고 공공기관 또한 투명하게 서비스를 관리할 수 있다. 또한 블록체인은 변경 불가능성(Immutability)을 제공하기 때문에, 데이터의 무결성에 있어 고령자나 대리인의 정보, 또는 일자리 매칭 과정에서 생성된 데이터는 기록된 후 임의로 변경될 수 없다. 이러한 특성 덕분에 시스템은 데이터의 신뢰성을 보장하며, 사용자는 시스템이 제공하는 서비스의 안정성을 확신할 수 있다.

참고문헌

[1] 진은하, 이철희, “블록체인 기술 및 보안 위협 분석”, 디지털산업정보학회논문지, Vol14, No.4, 47-56p, 2018.



(그림1) 블록체인 기반 아키텍처 개요

2.3 시스템 아키텍처

시스템의 아키텍처는 블록체인 기술과 AI 기반 일자리 추천 알고리즘이 결합된 형태로 설계되었다. 프론트엔드는 Flutter로 구현되어 다양한 디바이스에서 접근 가능하며, 서버는 Spring Boot와 MySQL을 통해 고성능으로 운영된다. 블록체인 기술은 네이버 클라우드 플랫폼(NCP) 상에 구현되며, 모든 데이터는 암호화된 상태로 저장되고, 분산 원장 구조를 통해 데이터의 무결성이 보장된다. 또한, PyTorch 및 TensorFlow를 활용하여 고령자 맞춤형 일자리 추천 시스템이 AI 기반으로 학습되며, 매칭된 일자리 데이터는 블록체인에 안전하게 기록된다.

3. 결론

본 논문에서는 블록체인과 인공지능 기술을 활용하여 고령자를 위한 맞춤형 일자리 매칭 시스템을 설계하였다. 블록체인의 높은 보안성과 투명성을 바탕으로 고령자들의 신뢰를 확보할 수 있었으며, 법정 대리인을 통한 인증 절차를 통해 사용자의 접근성을 높였다. 또한 AI 기반의 일자리 추천 시스템은 사용자의 관심사와 능력에 맞춘 적합한 일자리를 제공하여, 고령자들이 보다 쉽게 사회에 재진입할 수 있는 기반을 마련하였다.

인간 정보처리 특성을 고려한 안전 정보디자인: 감각, 지각, 인지, 기억단계별 디자인 전략

김다솔, 류시천¹⁾
조선대학교 대학원 창의공학디자인융합학과 박사과정,
조선대학교 디자인공학과
info-pedia@naver.com, scyou@chosun.ac.kr

Safety Information Design Based on Human Information Processing Characteristics: Design Strategies for Sensory, Perceptual, Cognitive, and Memory Stages

Kim Da-Sol, You Si-Cheon
Department of Design & Creative Engineering,
School of Informatics & Product Design, Chosun University, Gwangju, Korea

요 약

다양한 재해와 사고가 빈번하게 발생하고 인명 피해가 증가함에 따라, 중대재해처벌법의 처벌 규정이 강화되었다. 이러한 시대적 변화에 맞춰 디자인학 분야에서도 안전디자인의 중요성이 부각되었으며, 이에 대한 연구자들의 학술적 논의가 더욱 활발해질 필요가 있다. 본 연구의 목적은 안전 정보디자인의 관점에서 인간의 인적 특성을 반영한 정보처리 단계별 효과적인 정보 제시 전략을 탐색하는 데 있다. 문헌 조사를 통해 재난 상황에서 나타나는 인적 특성을 분석하고, 정보처리 단계별로 적합한 정보 제시 방안을 제안하였다. 주요 발견은 다음과 같다. 첫째, 감각 및 지각 단계에서는 스토리텔링, 비교·대조, 다중감각을 활용한 감각적 정보의 우선순위가, 인지 단계에서는 청킹, 맥락을 활용한 구조적 정보의 우선순위가 안전 정보디자인의 주요 지표로 작용할 수 있다. 둘째, 기억 단계에서는 기존의 시각적 자극에만 의존하던 정보 제시에서 벗어나, 후각을 포함한 다중감각 정보를 활용함으로써, 신속한 기억 인출을 촉진하고 안전한 상황에 도달하는 데 기여할 수 있다.

1. 연구의 필요성

위험한 상황에서 정보수용자는 이성적 판단을 상실하고 본능적인 행동을 취할 가능성이 높다. 이에 따라 위험 상황에서 발생하는 인간의 특성을 고려한 정보 제시 방안에 대한 학문적 논의가 요구된다. 본 연구의 목적은 안전 정보디자인 관점에서 위험 상황에서 나타나는 인간의 특성을 파악하고, 정보처리 단계별로 효과적인 정보 제시 전략을 탐색하는 것이다. 위험 상황에서 정보수용자의 정보처리 특징을 분석하고, 안전 정보디자인 관점에서 고려해야 할 인간 특성에 대한 이론적 고찰 내용을 제시하고자 한다. 또한, 관련 문헌을 통해 인간 특성을 반영한 정보처리 단계별 정보 제시 방안을 제안한다. 본 연구는 안전 정보 전달을 위한 가이드라인 마련에 기초가 되는 선행 연구로서, 위험 상황에서 발생하는 인간 특성과 이에 따른 정보 제시 방안에 대한 이론적 검토에 초점을 두었다.

2. 안전 정보디자인의 개념 및 특징

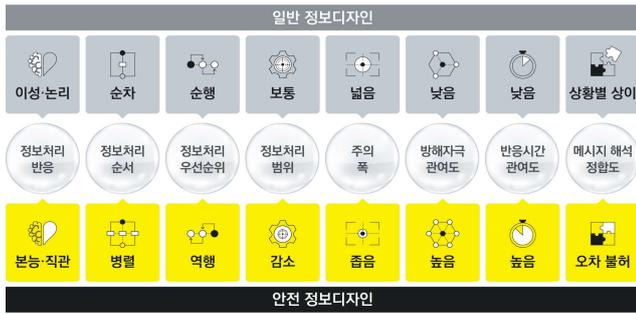
안전 정보디자인(Safety Information Design)은 위험 상황에서 최적화된 정보 제시 방안을 탐색하는 디자인 분야

이다. 심리학자 제임스 리즌(James Reason, 1984)은 위험 상황에서 발생하는 인간의 정보처리 오류로 인한 사고율을 연구한 결과, 감각 및 지각적 오류로 인한 사고율은 61%, 이해와 해석의 오류는 27%, 그리고 지식의 부정확성으로 인한 사고율은 11%임을 밝혔다[1]. 이를 통해 위험 상황에서의 정보처리 우선순위가 감각과 지각 > 인지 > 기억의 순서로 중요하다는 점을 알 수 있다.

3. 위험 상황에서 정보수용자의 정보처리 특징

위험 상황에 직면한 정보수용자는 본능적인 행동을 취할 가능성이 높다[2]. 감각 및 지각 단계에서는 눈에 잘 띄는 자극이 우선적으로 탐색되며, 이 과정에서 방해 자극이 많아도 병렬 처리로 신속하게 정보가 처리된다. 이러한 특성은 정보처리 범위와 주의 폭, 그리고 방해 자극의 관여도에 영향을 미친다. 인지 단계에서는 메시지의 이해, 판단, 반응이 안전 수준을 높이는 주요 요소로 작용한다. 기억 단계에서는 과거 기억과 새롭게 부호화된 정보가 결합되어 신속하고 정확한 반응을 유도할 수 있어야 한다.

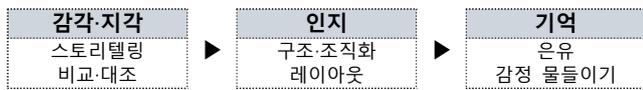
1) 교신저자 : 류시천, 조선대학교 디자인공학과 교수
e-mail : scyou@chosun.ac.kr



<그림 1> 안전 정보디자인의 특징

4. 위험 상황에서 정보처리 단계별 정보 제시 방안

류시천(2023)은 위험 상황에서 인간 특성을 반영한 다음의 정보 제시 방안이 안전성을 높이는 데 효과적이라고 주장하였다[3].



본 연구에서는 류시천의 이론을 바탕으로 문헌을 검토하고, 위험 상황에서 발현하는 인간 특성을 고려한 정보처리 단계별 정보 제시 방안을 분석하였다.

감각·지각 단계에서는 자극의 신속한 탐색을 유도하는 것이 중요하다. 현출성(salience)은 특정 자극이 다른 자극에 비해 더 쉽게 탐지되는 속성을 의미하며, 이는 시각뿐만 아니라 다양한 감각 탐지기에 적용된다. 스토리텔링은 정보를 생생하게 전달하고 정보수용자의 주의를 집중시키는 데 효과적이다. 또한, 비교는 서로 다른 대상의 공통점을, 대조는 유사한 대상의 차이점을 드러내는 방식으로[4], 두 가지 이상의 요소를 비교하여 정보의 패턴을 인식하고 중요한 차이를 감지할 수 있게 한다. 다양한 감각을 자극하는 정보 제공은 감각·지각 처리의 신속성을 돕는 데 유용하다.

인지 단계에서는 정보의 구조화와 조직화, 레이아웃 설계가 중요한 역할을 한다. 이 과정에서 청킹(Chunking) 기법이 유용하다. 청킹은 관련 있는 정보를 묶어 의미 있는 단위로 만들고, 이를 통해 정보의 빠른 이해를 돕는다. 또한, 청킹은 게슈탈트(Gestalt) 지각이론, 아른하임(Rudolf Arnheim)의 균형이론, 깁슨(Gibson)의 생태학적 지각이론에 근거하여 설명될 수 있다[5]. 게슈탈트 이론은 형상과 배경 간의 관계를, 아른하임 이론은 구성 요소 간 힘의 균형을, 깁슨 이론은 시각적 배열의 유형을 설명하며 청킹의 효과를 뒷받침한다. 이와 함께, 맥락을 제공하는 정보는 정보수용자의 연상 작용을 촉진하여 신속한 판단을 돕는다. 예를 들어, ‘화재 경고’ 메시지와 함께 관련 맥락 이미지를 제공하면 정보수용자는 화재 위험을 즉시 인식하고 반응할 수 있다.

기억 단계에서는 과거 기억의 회상과 신속한 인출을 유도하는 것이 중요하다. 감정이 담긴 정보는 장기 기억으로 처리될 가능성이 크다. 강렬한 감정적 경험은 해마와 편도체의 활성화를 촉진하여 기억 형성 능력을 높인다. 특히,

감각과 기억의 결속은 기억을 부호화하고 인출하는 데 효과적이다. 예를 들어, 후각과 연결된 기억은 다른 감각에 비해 오랜 기간 유지된다는 연구 결과가 있다[6]. 또한, 은유는 복잡하거나 추상적인 개념을 친숙한 개념으로 연결하여 설명함으로써 기억을 강화하고 필요한 순간에 쉽게 회상할 수 있도록 돕는다. 은유는 위험 상황의 심각성을 강조하고, 특정 행동의 중요성을 더 강력하게 전달하는 데 유용하다.

단계	감각·지각	인지	기억
핵심지표	정보우선순위		기억스키마 형성
	[감각적]...현출성	[구조적]...특징탐색	장기기억화
	기억 인출		기억 부호화
정보 제시방안	스토리텔링	청킹	은유
	비교·대조	맥락	감정 물들이기
	다중감각	다중감각	다중감각

<그림 2>. 인적 특성을 고려한 정보 제시 방안

5. 결론

본 연구는 안전 정보디자인 관점에서 정보처리 단계별 효과적인 정보 제시 방안을 탐색하였다. 감각·지각 단계에서는 자극의 현출성을 높이는 스토리텔링, 비교·대조, 다중감각 제공이 효과적임을 발견했다. 인지 단계에서는 청킹과 구조화된 맥락 제공이 이해와 판단을 돕는 주요 요소로 작용하며, 기억 단계에서는 은유와 감정물들이기, 다중감각 활용이 위험 상황에서 인간의 기억 인출과 부호화를 촉진하는 데 유효한 전략임을 확인하였다.

참고문헌

[1] 이순열, 안전심리학, 학지사, 2019, pp.86-96
 [2] Ntokos, Konstantinos, Level of Fear: Analysis of Fear Spectrum into a tool to support Horror Game Design for Immersion and Fear, 2018, p4
 [3] 류시천, ‘지식시각화’ 교과목 수업자료, 조선대학교 대학원 창의공학디자인융합학과, 2023, pp.11-12
 [4] E. Bruce Goldstein, Laura Cacciamani, 감각 및 지각 심리학(Sensation & Perception), 박학사, 2023, p.156
 [5] 윤성원, 디자인 균형의 원리 시각적 복잡성, 미세음, 2020, p.91
 [6] E. Bruce Goldstein, Laura Cacciamani, op. cit., pp. 485-486.

신입생 및 복학생의 학교 적응을 위한 소모임 어플리케이션*

이규민, 김종윤, 이민우, 심훈, 전태민, 송지영
 한남대학교 컴퓨터공학과
 e-mail : kmin3625@gmail.com, jysong@hnu.kr

Small Group Application for New and Returning Students' Adaptation to School

Lee Gyu-min, Kim Jong-yun, Lee Min-woo, Sim Hoon, Jeon Tae-min Jiyoung Song
 Department of Computer Engineering, Hannam University

요 약

신입생과 복학생들은 대인관계 형성을 어려워하는 문제가 있다. 이러한 문제를 겪는 사람들은 소모임에 참여하는 것이 도움이 되지만, 학생들이 소모임을 만들고 친구를 사귄 수 있는 도구나 수단이 부족한 것이 현실이다. 그래서 학교생활에서 소외되지 않도록 도와주고, 자신의 관심사나 전공을 다른 사람들과 공유하고 발전할 수 있는 앱 ‘뽕’을 제안한다. 본 연구에서는 학교 안에서 진행되는 소모임 앱으로 문제를 해결한다. 이 앱에는 크게 두 가지 주기능이 있다. 첫 번째, 학교 지도에 목적별 모임을 마커로 표시해 실시간 소모임을 만드는 것이다. 두 번째, 만들어진 소모임에서 사용자의 소통을 위해 채팅 기능을 지원하는 것이다. 이를 통해 소모임을 쉽게 만들고 참여하여 다양한 사람들과 어울리고 교내에 활력을 줄 것으로 기대한다.

1. 서 론

신입생들이나 복학생들은 학교생활에 적응하는 데 어려워하는 문제가 있다. 이러한 문제들은 학교를 자퇴하거나 제적당하는 등의 대학 중도 탈락으로 이어지기도 한다. 한수미 외[1]에 의하면 중도 탈락률은 20년도 4.3%, 21년도 4.6%, 22년도 4.9%로 대학 중도 탈락률은 매년 꾸준히 늘어나고 있다. 대학 중도 탈락률은 대학에서 심각한 문제로 대두되고 있고 대학 적응 실패 등을 원인으로 꼽고 있다. 대학 적응 실패를 막을 수 있는 대안으로 대학 온라인 커뮤니티가 있다. 황보경 외[2]에 의하면 코로나 이후 20대는 44.3%가 대학 커뮤니티 어플리케이션을 이용한다고 한다. 이처럼 코로나 이후 대학 커뮤니티 앱을 사용하는 사람들이 늘어나고 있지만 직접 만나 취미 공통관심사를 공유하거나 식사를 같이하는 등 소모임을 만들고 친구를 사귄 수 있는 어플리케이션이나 커뮤니티가 부족하다. 본 연구에서는 학생들 사이 소모임을 만들어 상호작용하기 좋은 환경을 만들고 만남을 도와주는 앱인 ‘뽕’을 제안한다.

2. 관련 연구

2.1 소셜 매칭 방식

안진호 외 [3]의 논문은 소모임 매칭 시스템으로 두 가지를 제안한다. 첫 번째 실시간 매칭 시스템이란 해당 조건에 맞는 가장 가까운 사람이 감지되면 상대방의 정보를 제공하는 창이 나오고 이를 수락 또는 거절할 수 있다. 양측 모두 수락하여 매칭이 이루어진다면 채팅이 가능해지고 지도에 핀을 이용하여 각자의 위치를 표기해준다. 두

번째, 소셜 매칭 시스템은 게시자가 글을 작성함으로써 요청이 들어오면 게시자의 위치 정보를 수집하여 유저 위치 정보 데이터베이스에 보관 후 참가자가 게시글을 활성화하는 경우 해당 정보를 제공한다. 마찬가지로 참가자가 게시글을 활성화하는 경우 유저 개인 정보 데이터베이스에서 참가자 정보를 제공하여 해당 게시글의 조건과 유저 정보가 부합하는지 확인한다. 참가자가 소모임에 참여하게 된다면 채팅 서비스로 연결된다. 본 논문은 소셜 매칭 시스템을 채택하여 연구를 진행한다.

2.2 실시간 채팅 시스템 구현을 위한 소켓 기반 방식

고완기 외 [4]의 논문에서는 소켓 기반 모바일 SNS 연동 시스템을 설계 및 구현하여 클라이언트 간의 실시간 통신을 가능하게 한다. 본 논문에서는 이 소켓 기반 방식을 참고하여 실시간 채팅 시스템을 구현하고자 한다. 소켓 기반 채팅 방식은 클라이언트와 서버 간의 신속하고 안정적인 데이터 송수신을 가능하게 하며, 특히 모바일 환경에서 효과적이다. 이를 통해 사용자 간의 실시간 채팅이 원활하게 이루어질 수 있도록 시스템을 설계한다.

2.3 기 구현된 어플리케이션 및 서비스와 비교 평가

특성 어플	소모 임	지도 기반	익명성	채팅	학생 대상
포켓유니브	X	X	O	O	O
턴더	X	O	O	O	X
당근	O	O	O	O	X
에브리타임	X	X	O	O	O
뽕	O	O	O	O	O

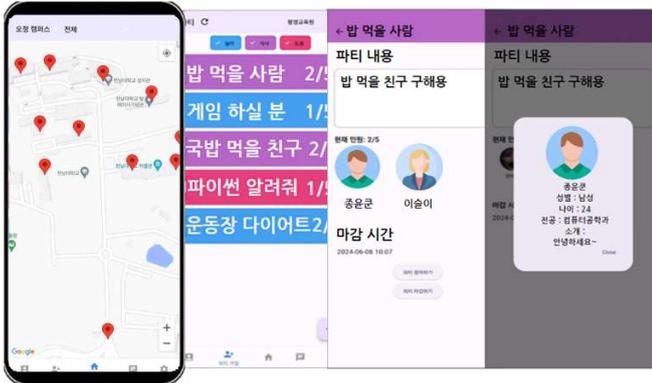
[표 1] ‘뽕’과 기 구현된 서비스와의 비교

* 이 논문은 2024년도 한남대학교 지방대학활성화 사업의 지원을 받아 수행된 연구임

표 1은 소모임 어플리케이션이 가져야 할 특성을 선정 한 후 기 구현된 유사 어플리케이션 및 서비스와의 비교를 평가로 나타낸 것이다. 학교 관련 앱인 포켓유니브, 에브리타임은 학생 대상, 익명성, 채팅 특성을 만족하지만, 소모임과 지도 기반의 특성을 만족하지 못한다. 당근과 틴더는 지도 기반이지만 학생 대상이라는 특성을 만족하지 못하였다. 또한, 틴더는 소모임 특성을 만족하지 못한다. ‘뽕’은 선정한 모든 앱 특성을 만족하여 학생들이 교내에서 안전하고 편리하게 소모임을 진행할 수 있도록 돕는다.

3. 지도 기반 소모임 어플리케이션 설계 및 구현

3.1 구글 맵 API를 활용한 메인 화면과 소모임 게시판

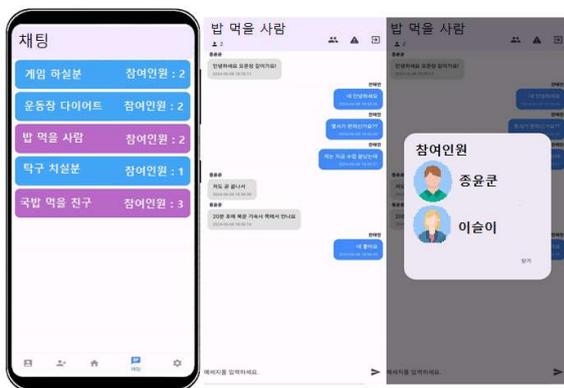


[그림 1] 메인 지도 화면과 소모임 게시판

본 연구는 구글 맵 API를 활용해 지도 기반 메인 화면을 구성하였다. 지도의 건물들에는 각 게시판과 연결된 마커를 표시하였다. 지도 기능은 부가 기능인 현 위치를 알려주는 기능과 함께 상승효과를 만들어 학교에 새로 적응하는 사람들에게 직관적으로 어떤 장소에서 소모임이 진행되는지 알려주는 효과적인 도구로 활용된다.

각 마커를 눌러 연결되는 각 게시판에는 3가지 모집 종류인 ‘놀이’, ‘식사’, ‘도움’으로 분류된 소모임 모집 글이 있다. 사용자는 실시간으로 소모임을 만들거나 참여할 수도 있으며, 게시판의 소모임을 눌러 상세 정보를 보면 구체적인 내용과 참여 인원의 정보를 확인할 수 있다.

3.2 소통과 협력을 위한 실시간 채팅

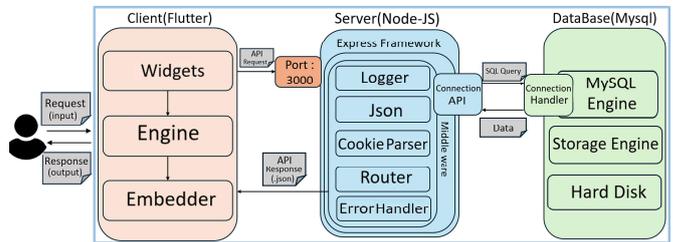


[그림 2] 채팅 화면

그림2는 채팅 화면이다. 채팅 기능은 소모임 활동을 원활하게 진행하도록 돕는 중요한 요소이다. 실시간 메시지 전

송과 대화 기록 저장을 통해서 모임 일정 조율과 정보를 빠르게 공유하게 한다. 또한, 부적절한 내용 신고와 소모임 인원 정보 확인 기능을 제공하여 안전한 소통 환경을 유지한다.

3.3 서버, 클라이언트, DB 상호작용을 위한 시스템 설계



[그림 3] 프레임워크

본 연구는 Flutter, Node.js, MySQL을 활용하여 설계하였다. 그림3은 User의 요청을 통해 Client, Server, DB가 어떻게 상호작용을 하는지 보여준다. User의 요청을 받을 시 Client에서는 API Request를 Server에 보낸다. Server에서는 받은 요청을 미들웨어에서 처리한 후 라우팅하여 해당 요청을 처리할 핸들러 함수를 호출한다. 호출된 핸들러 함수는 요청된 데이터를 조회, 조작하기 위해 DB에 SQL Query문을 보낸다. DB에서는 Query문을 처리하고 데이터를 반환한다. Server는 DB로부터 받은 데이터로 응답 데이터를 생성하여 Client에 반환한다. Client는 Server로부터 받은 응답 데이터를 처리해 User에게 결과를 제공한다.

4. 결론

본 연구에서는 신입생이나 복학생 같은 학교생활에 적응하기 어려운 사람을 위한 지도 기반 소모임 매칭 메신저 앱 ‘뽕’을 제안한다. ‘뽕’은 신입생과 복학생을 포함한 대학생들이 더욱 쉽게 새로운 대인관계를 형성하고 캠퍼스 생활에 적응할 수 있도록 돕는 도구가 될 것이다. 학교 지도를 기반으로 한 모임 생성 기능과 실시간 채팅 기능을 통해 학생들이 자신의 관심사나 전공을 공유하고 소통할 기회를 제공한다. 결과적으로 활기찬 캠퍼스 문화를 조성하는 데 이바지할 것으로 기대된다.

참고문헌

- [1] 한수미, 권병국 / 2023 / 코로나 팬데믹 이후 대학생의 대학생활 만족도, 학업성취도, 자아존중감이 대학생활적응에 미치는 영향 / 인문사회 21, vol. 14, no. 3, 통권 58호, pp. 1621-1636
- [2] 황보경, 김승인 / 2022 / 대학 생활 커뮤니티 어플리케이션의 사용자 경험 연구 / 차세대컨버전스정보서비스기술논문지, vol. 11, no. 6, 통권 34호, pp. 655-665
- [3] 안진호, 송재민 / 2023 / 식사 메이트 어플리케이션, ‘밥풀’ / Proceedings of KIIT Conference, vol. 2023, no. 6, pp. 860-864
- [4] 고완기, 고석용, 김제석 / 2013 / SOCKET기반 모바일 SNS 연동시스템 설계 및 구현 / 디지털융복합연구, vol. 11, no. 8, pp. 367-373

안면 색상 데이터 기반 퍼스널 컬러 진단 서비스*

김주희, 김채하, 이가은, 송지영

한남대학교 컴퓨터공학과

lsj240606@naver.com, kimcheaha0617@gmail.com, dlrkdms001@gmail.com, jysong@hnu.kr

Facial Color Data-based Personal Color Diagnosis Service

Juhee Kim, Chaeha Kim, Gaeon Lee

Department of Computer Engineering, Hannam University

요 약

코로나19 팬데믹이 종료되면서 의류 소비량이 증가하는 추세이다. 이에 따라 의류에 대한 관심이 증가하고 있으며, 개인이 가진 신체의 색과 어울리는 색을 의미하는 퍼스널 컬러에 대한 관심도 커지고 있다. 많은 사람들이 자신의 퍼스널 컬러에 맞추어 옷을 구매하는 추세가 강조되고 있으나, 퍼스널 컬러 진단 비용은 1인 평균 10만 원으로 많은 소비자가 부담을 느끼고 있다. 본 논문에서는 안면 색상 데이터를 기반으로 한 퍼스널 컬러 진단 서비스를 제안함으로써 위 문제를 극복한다. 해당 서비스의 기능은 이미지에서 얼굴을 인식하여 얻은 안면 색상 데이터를 기반으로 퍼스널 컬러를 진단한다. 해당 서비스를 통하여 서비스 이용자들이 자신의 퍼스널 컬러 톤을 인지하여, 실생활에서 퍼스널 컬러에 맞는 코디를 착용할 수 있도록 돕고자 한다.

1. 서 론

코로나19 팬데믹이 종료되면서 의류 소비가 급증하고 있다.[1] 이는 사람들의 패션에 대한 관심을 더욱 고조시키고 있으며, 특히 퍼스널 컬러에 대한 수요가 급증하고 있다. 퍼스널 컬러는 개인의 피부, 머리카락, 눈동자 색상에 기반하여 가장 잘 어울리는 색상을 찾는 과정으로, 이를 통해 개인의 스타일링 만족도를 향상할 수 있다.[2] 그러나 1인 평균 10만 원에 달하는 퍼스널 컬러 진단의 높은 비용은 많은 소비자들에게 큰 부담으로 작용하고 있다.

현재 시장에 존재하는 온라인 퍼스널 컬러 진단 서비스들은 전문가의 판단 없이 진행되는 일종의 자가 진단 서비스로, 진단 과정에 주관적인 요소가 개입될 수 있어 결과의 신뢰성이 떨어지는 문제가 있는 것으로 파악됐다. 이러한 문제를 해결하기 위해 본 논문에서는 안면 색상 데이터 분석을 기반으로 한 퍼스널 컬러 진단 서비스를 제안한다. 본 서비스는 AI 기술을 활용하여 진단 결과의 신뢰성을 높이며, 실시간 접근이 가능한 온라인 플랫폼으로써 이용자들이 시간과 비용의 제약 없이 자신의 퍼스널 컬러를 진단받을 수 있도록 한다.

2. 관련 연구

퍼스널 컬러는 봄, 여름, 가을, 겨울의 4계절 유형으로 분류된다[3]. 각각의 계절 유형에 맞는 색상 팔레트가 제안된다. 따뜻한 느낌을 주는 워몬은 노란빛이 도는 톤이며, 차가운 느낌을 주는 쿨톤은 푸른빛이 도는 톤이다.[4]

㈜작당모의에서 개발한 쟁페이스[5]는 가상 메이크업 기술을 구현하여 사용자가 자신에게 어울리는 퍼스널 컬러를 직접 고르는 자가 진단 서비스를 제공한다. 비슷한 사례로, 색간다 팀에서 개발한 오빠 톤 많아?[6] 웹 서비스는 사용자가 자신의 이미지를 업로드하면, 퍼스널 컬러 별로 다양한 톤의 색채 이미지를 배경으로 한 사용자의 얼굴 사진을 제공한다. 사용자는 이 배경 이미지를 보고 잘 어울리는 색상을 직접 골라 자가 진단한다. 기존 연구에서는 사용자가 자가 진단하는 과정에서 개인의 주관에 개입하여 정확한 결과를 얻기 어려운 문제점이 있다. 이러한 문제를 해결하기 위해 본 연구에서는 피부색의 노란빛 차이를 기준으로 퍼스널 컬러를 4계절 유형으로 나누어 진단하는 AI 기술을 제안한다. 해당 기술은 사용자 개인의 주관이 개입되지 않아 보다 정확한 결과를 제공한다.

3. ColorFinder

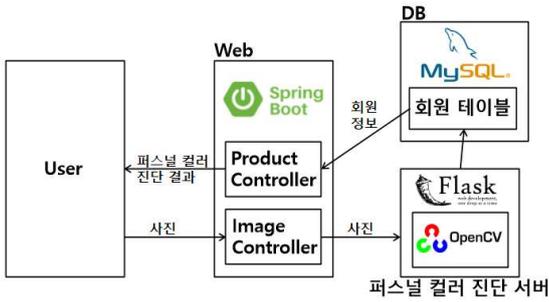
3.1 서비스 설계

서비스 내부 구현 부는 크게 3가지로 나뉜다. 각각은 웹 서버, 데이터베이스, 사용자 퍼스널 컬러 진단 서버로 구분된다. 사용자가 퍼스널 컬러 진단을 위해 웹서버에 이미지 파일을 전송하면 웹서버의 ImageComtoller가 사진을 사용자 퍼스널 컬러 진단 서버인 flask서버에 전달한다. 전달된 사진은 OpenCV 라이브러리를 사용하여 만들어진 파이썬 코드에 의하여 퍼스널컬러가 진단된다. 진단 결과는 데이터베이스의 회원 테이블에 저장되며, 해당 정보는 다시 웹서버의 ColorViewControllor에 의해 사용자에게 반환된다.

3.2 퍼스널 컬러 진단

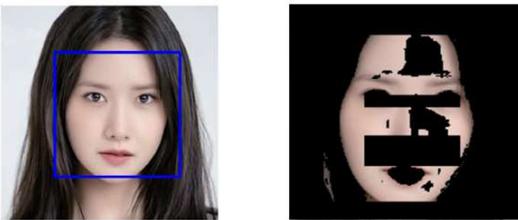
* 이 논문은 2024년도 한남대학교 지방대학활성화 사업의 지원을 받아 수행된 연구임

진단 첫 단계에서 사용하는 얼굴 인식 기술은 인텔이



(그림 1) 서비스 프레임워크

개발한 OpenCV 라이브러리를 사용한다. 얼굴 인식 자동화 함수를 호출하여 검출된 부분으로 얼굴 범위를 지정한다. 지정된 얼굴 범위에서 inRange() 함수를 적용하여 사람의 피부색에 해당하는 상한값과 하한값을 지정한 뒤, 해당 범위 내의 픽셀을 추출한다. Jorge Alberto 외[7]의 연구에 따르면, 사람의 피부는 YCbCr 색공간 내의 (0,133,77)~(255,173,127) 영역 내에 존재한다. 이에 따라 피부색 범위를 지정한다. 그림 2는 이미지의 얼굴 인식과 인식된 범위에서의 추출된 피부색을 나타낸다.



(그림 2) 이미지의 얼굴 인식과 추출된 피부색

추출된 피부색 값을 LAB 색 공간의 B 채널 값으로 변환한다. 해당 값은 파란색과 노란색의 차이를 수치로 나타낸 값이다. 이 값이 클수록 피부색이 노란빛을 많이 띠고, 적을수록 파란빛을 많이 띤다는 의미이다. 이 차이를 이용하여 웜톤과 쿨톤을 구분한다.

4계절의 퍼스널 컬러 별로 예시 연예인의 사진을 100장씩 수집한 뒤, 동일한 진단 과정을 진행한다. 계산된 B 채널 값의 평균을 내어 퍼스널 컬러를 분류할 수 있는 표본 데이터셋을 구축한다. 이 데이터셋과 사용자 이미지의 B 채널 값을 비교하여, 가장 유사한 값의 퍼스널 컬러를 진단 결과로 반환한다. 퍼스널 컬러 진단 결과는 그림 3과 같이 색상 팔레트와 함께 제공된다.



(그림 3) 퍼스널 컬러 진단 페이지

4. 실험 결과

서비스의 정확도를 측정하기 위해 전문 컬러리스트 박예영[8]의 진단 결과를 기준으로 퍼스널 컬러 유형별 인물 사진을 5장씩 수집하여 퍼스널 컬러 유형별 진단 정확도 비교 실험을 진행하였다. 표 1은 정확도 측정 결과를 도표로 정리한 그림이다. 정확도는 유형별 실험 중에서 일치한 실험 결과의 비율을 백분율로 나타낸 값이다. 진단 결과는 컬러별 편차가 존재하나, 60 ~ 100%로 높은 정확도를 가지고 있다.

(표 1) 퍼스널 컬러 진단 서비스 정확도

		1	2	3	4	5	정확도
컬 톤	여름						60%
	겨울						80%
웜 톤	봄						100%
	가을						100%

5. 결 론

본 연구에서는 안면 색상 데이터를 기반으로 한 퍼스널 컬러 진단 및 맞춤형 의류 추천 서비스를 제안한다. 이미지에서 얼굴을 인식하여 안면 색상을 추출하고, 퍼스널 컬러를 진단하는 알고리즘을 제공한다. 이에 사용자들은 자신의 퍼스널 컬러에 맞는 의류를 구매하여, 의류 소비 만족도를 높일 수 있다.

참 고 문 헌

- [1] 통계청, 「4/4분기 가계동향조사 결과」, 2023, 1024.06.16., 전국가구 가계지출.
- [2] 고아라, 이영재. “퍼스널컬러가 유행인식, 자기이미지에 미치는 영향”, 한양대학교, 디자인학과, 2010, pp.46-62.
- [3] 임도연, “퍼스널 컬러와 토탈 코디네이션 연출에 관한 연구”, 콘텐츠와 산업 제5권, 2023
- [4] Jackson, C. (1987), Color Me Beautiful. NY: Ballantine Books.
- [5] 작당모의, “잼페이스” [모바일 어플리케이션], 2023
- [6] 색간다, “오빠 톤 많아?” [모바일 웹사이트], 2023
- [7] Jorge Alberto Marcial Basilio, “Explicit image detection using YCbCr space color model as skin detection”, 2011
- [8] 퍼스널컬러 컨설팅 전문기업 ‘PL COLOER’ 대표, 퍼스널컬러 컨설턴트 자격 심사위원, 전문 컬러리스트

초음파 센서 기반 실시간 빗물받이 관리 시스템*

배정훈, 최지영, 최지민, 국지언, 장민성, 송지영
 한남대학교 컴퓨터공학과
 e-mail : qowjdgns0106@naver.com, jysong@hnu.kr

Ultrasound Sensor-Based Real-Time Rainwater Drainage Management System

Bae Jung-hun, Choi Ji-young, Choi Ji-min, Kuk Ji Eon,
 Jang Min-sung, Jiyoung Song
 Department of Computer Engineering, Hannam University

요 약

지구온난화의 영향으로 집중호우가 빈번히 발생하고 있다. 집중호우의 빈번한 발생은 침수의 원인이 되고 재산, 인명 피해가 발생한다. 침수의 주원인 중 하나는 배수 기능의 상실이다. 도로에는 빗물을 배수관으로 흘려보내기 위해 빗물받이가 설치되어 있다. 빗물받이에 쓰레기, 낙엽 등으로 인해 이물질이 쌓이게 되면 빗물을 흘려보내지 못하여 침수가 발생한다. 빗물받이를 관리하기 위한 인력과 비용에 비해 빗물받이의 개수가 많아 효율적인 관리가 어렵다는 문제가 있다. 본 논문에서는 초음파 센서를 이용한 빗물받이 관리 시스템을 제안함으로써 위 문제를 극복하고자 한다. 첫째, 초음파 센서를 활용해 빗물받이의 내부 용량을 확인한다. 용량이 설정된 기준을 초과하는지의 여부를 초음파 센서를 통해 확인한다. 둘째, 웹 페이지를 통해 빗물받이의 전체적인 관리를 한다. 기준을 초과하는 빗물받이의 경우 색상 변화를 통해 즉각적으로 확인이 가능하도록 설계한다. 초음파 센서를 통해 빗물받이의 상태를 확인하고, 관리함으로써 적은 비용으로 침수를 예방할 수 있을 것으로 기대한다.

1. 서 론

세계기상기구 발표에 의하면, 2021년 지구 평균기온은 산업화 이전 수준보다 약 1.11°C 높아 2015년 이후 7년 중 가장 따뜻한 한 해로 기록되었다. 최근 이상기후로 인해 짧은 시간 동안 많은 비가 내리는 집중호우가 증가하고 있다. 2022년 8월 서울에는 시간당 130mm, 하루 360mm의 집중호우가 발생했다. 짧은 시간 비가 집중되면서 물이 역류해 침수되는 경우가 빈번히 일어나고 있다. 국토연구원[6]의 문서에서는 침수 원인의 27퍼센트는 빗물받이의 이물질로 인한 배수 능력 저하가 문제라고 발표하였다. 현재 빗물받이의 문제점은 관리 필요 여부를 직접 확인해야 한다는 것이다. 자동화된 측정 도구가 없어 도시의 모든 빗물받이를 관리하기에는 비용이나 인력에 한계가 존재한다. 침수를 예방하기 위한 빗물받이 관리 방법이 부족한 것이 현실이다. 본 연구에서는 빗물받이의 이물질을 자동으로 체크하고 관리자가 실시간으로 확인할 수 있는 방법을 제시하고자 한다.

2. 관련 연구

2.1 침수 취약 예측

도시지역에서 발생한 침수의 다수는 빗물받이를 통해 배수되

지 못한 물과 배수량을 상회하는 물의 유입으로 인한 것이다. 손아롱 외[1]의 논문은 과거 침수 자료를 기반으로 도시의 침수 지역을 예측하는 방법론을 제시한다. 침수 데이터를 통한 정확한 예측이 가능하지만, 국소 범위만을 측정 가능하다. 김광영 외[2]와 송영미 외[3]는 CCTV, 침수심 센서, SNS 등 데이터를 활용해 실시간 대응 시스템을 제시한다. 실시간으로 광범위한 측정은 가능하지만, 사용자 데이터에 의존해야 한다는 단점이 존재한다. 본 논문에서는 기존 데이터에 의존하지 않고, 실시간 측정값을 통해 침수를 예방할 수 있는 시스템 연구를 진행한다.

2.2 빗물받이 관리

빗물받이에 이물질이 쌓이면 물의 흐름이 막혀 침수가 발생하게 된다. 침수 예방을 위해서는 빗물받이의 지속적인 관리가 중요하다. 광명시[4]의 기사에서는 빗물받이에 스티커를 부착해 관리하는 방법을 제시한다. 해당 관리 스티커를 통해 앱으로 신고하면 청소를 위한 인력이 배치된다. 비용이 적게 든다는 장점이 있지만, 주민 신고에 의존해야 한다. 진학성 외[5]의 문서에서는 비 유무에 따라 빗물받이 덮개가 자동으로 여닫는 시스템을 제안하였다. 해당 시스템은 물을 감지해 자동으로 동작하기 때문에 신고에 의존하던 관리 시스템에 비해 진일보한 방식이다. 다만 이물질이 덮개 위에 쌓여 있는 경우, 시스템 동작을 방해하여 침수를 유발할 수 있다는 문제점을 내포하고 있다. 본 논문에서는 초음파 센서를 포함한 소형 장치로 빗물받이 내의 이물질을 측정해 관리를 돕는 관제 시스템을 연구한다.

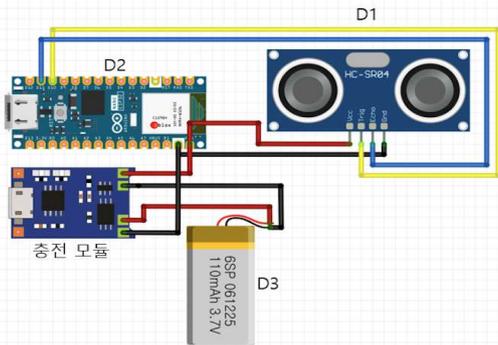
* 이 논문은 2024년도 한남대학교 지방대학활성화 사업의 지원을 받아 수행된 연구임

3. 빗물받이의 이물질 데이터 측정 및 관리

기존 빗물받이 관리 시스템은 물리적 차단이나 주민들의 신고에 의존하는 수동적인 시스템이었다. 본 연구는 초음파 센서를 활용하여 이물질의 높이를 실시간으로 측정하고, 측정 데이터를 웹 페이지를 통해 관리하는 통합적인 솔루션을 제안한다.

3.1 이물질 높이 측정을 위한 초음파 장치

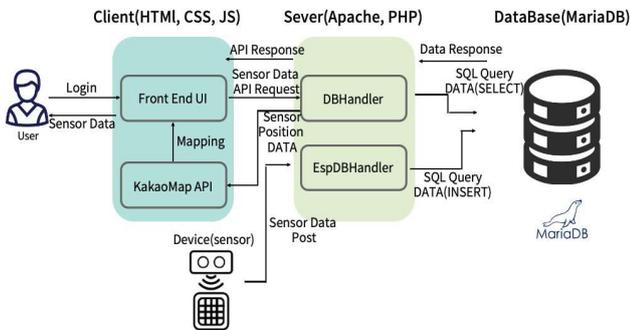
본 연구에서는 이물질 체크를 위한 센서로 초음파 센서를 채택하였다. 이물질과 물의 높이를 동시에 파악할 수 있는 초음파 센서는 집중 호우시 배수 용량 초과로 인한 침수를 조기에 감지하여 적시에 대응할 수 있다.



(그림 1 장치 구성도)

그림 1은 빗물받이 뚜껑에 설치할 장치의 배선도이다. D1은 초음파 센서로 이물질을 감지해 남은 용량 정보를 제공하며, D2의 Nano ESP32 모듈은 Wi-Fi를 통해 데이터를 서버로 전송한다.

3.2 실시간 관제를 위한 웹 페이지 구성



(그림 2 프레임워크)

그림 2는 전체 프로세스를 도식화한 것이다. 장치는 빗물받이 내부의 용량을 측정한 데이터를 EspDBHandler를 통해 데이터베이스에 저장한다. 서버는 DBHandler를 통해 측정 데이터를 실시간으로 조회한다. 측정 데이터는 웹 페이지에서 시각적으로 표현되어 관리자가 쉽게 확인할 수 있으며, KakaoMapAPI를 활용하여 빗물받이의 위치 정보를 지도에 표시한다.

그림 3은 웹 페이지의 구성을 보여준다. 지반이 낮은 곳 혹은 하천 주변은 침수가 빨리 일어날 수 있고 빗물받이 마다 깊이나 크기가 다르다. 관리자는 빗물받이 주변 상황에 맞게 빗물받이 내부 한계 용량을 미리 설정해둘 수 있다. 미리 설정해둔 빗물받이 내부 용량을 초과하면 PMD-S001과 같이 빨간색 테두리로 위험 경고를 표시한다. 장치의 활성화 여부는 PMD-S002와 PMD-S003과 같

이 장치의 이름 오른쪽에 초록색과 빨간색으로 표시된다.



(그림 3 웹 페이지 구성)

침수가 예상될 경우, 시스템은 해당 지역에 설치된 초음파 센서들이 동시다발적으로 경고 메시지를 보내게 된다. 경고 메시지를 통해 관리자는 침수 발생 구역을 예측하고, 신속한 대응 조치를 취하여 침수 피해를 예방할 수 있다.

4. 결론

본 연구는 침수의 원인 중 하나인 빗물받이 관리를 위한 시스템을 개발한다. 초음파 센서를 통해 각 빗물받이의 상태를 실시간으로 확인함으로써 관리자의 원활한 인력 배치를 도와준다. 본 연구의 시스템을 사용함으로써 적은 인력과 비용으로 빗물받이 관리 효율성 증가를 기대한다. 빗물받이의 관리 효율이 증가함에 따라 침수 피해가 감소할 것으로 판단된다. 본 연구는 실시간 측정 데이터를 통한 빗물받이의 효율적 관리를 지원함으로써 기존의 수동적인 시스템의 한계를 보완하는 데에 의의가 있다.

참고 문헌

- [1] 손아롱, 김병현, 한건연, "도시지역 도로 네트워크를 활용한 침수지역 예측에 관한 연구," 한국토목학회논문집, 제35권, 제2호, pp.307-318, 2015. doi: 10.12652/Ksce.2015.35.2.0307
- [2] 김광영, 양명석, "도시 침수 예측에서 대응까지: 솔루션 핵심기술 및 활용 사례 중심," KISTY 이슈 브리프, 제 61호, pp. 1-20, 2023. doi: 10.22810/2023KIB061.
- [3] 송영미, 김창수, "도로 침수 영역의 탐색을 위한 빅데이터 분석 시스템 연구," 한국멀티미디어학회 논문지, 제18권, 제8호, pp. 925-934, 2015. doi: 10.9717/kmms.2015.18.8.925.
- [4] 광명시, "광명시, 빗물받이에 관리번호 부착...도로침수 시 신속 복구," 광명시뉴스포털, 2021. [온라인]. 사용 가능: https://news.gm.go.kr/bbs/view.html?idxno=342&sc_category= [접속일: 2024년 6월 11일].
- [5] 진학성, 김창신, 이동섭, "스마트 빗물받이 관리 시스템," 서울시 치수과, 2022.
- [6] 국토연구원, "우리나라 도시침수피해 특성과 정책과제," 방재연구, vol. 8, no. 1, pp. 1-25, Jul. 2008.

클라우드 백엔드 비용 최적화를 위한 클라이언트 캐싱 기법 설계 및 구현

서지용, 장준혁
한남대학교

e-mail : officialcaox@gmail.com, jhjang@gmail.com

Cloud Cost Optimization through Client-Side User Data Caching Scheme

jiwoong Seo, Joonhyouk Jang
Hannam University

요 약

다양한 소프트웨어에서 클라우드 기반 서비스를 활용하는 사례가 증가하면서, 이에 따른 인프라 부담과 운영 비용이 함께 상승하고 있다. 이는 클라이언트의 속도 저하와 백엔드 서비스 이용자의 요금 부담을 가중시키고 있다. 이러한 문제를 해결하기 위해, 제안 시스템은 서버 요청 빈도를 줄이고, 이를 통해 서버 부하를 경감시켜 클라우드 운영 비용을 절감하는 것을 목표로 클라이언트 측에서 데이터를 캐싱한다. 또한, 캐싱된 데이터는 암호화를 통해 보안성을 높이며, 데이터 무결성을 유지하기 위한 데이터 처리 과정을 설계했다. 제안 시스템을 상용 소프트웨어의 프로토타입으로 구현하였다. 본 연구에서는 이러한 시스템의 설계와 구현 과정을 상세히 다루고, 실험을 통해 성능 향상과 비용 절감 효과를 검증하였다.

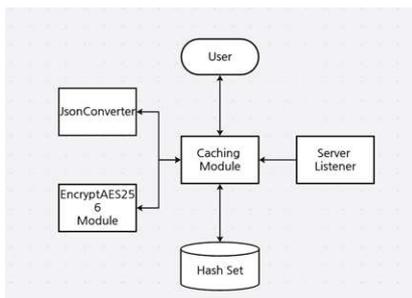
1. 서 론

클라우드 데이터베이스 컴퓨팅 요금제는 보통 요청, 데이터 저장, 데이터 전송 등의 활동에 따라 요금을 매긴다 [1]. 이로 인해 잦은 서버 요청은 네트워크 접속 속도 저하와 함께 비용 증가를 초래할 수 있다. 예를 들어, 시스템 설계 상의 결함 또는 구현 코드의 부주의로 인해 지나치게 많은 요청이 서버로 전송될 수 있고, 결과적으로 불필요한 비용이 과도하게 청구될 수 있다.

따라서, 이 논문에서는 클라이언트 측에서 사용자 데이터를 캐싱하는 기법을 제안한다. 또한 보안을 위해 사용자 데이터를 암호화하여 보관하도록 하였다. 제안 기법을 통해 서버 요청을 감소시킴으로써 운영 비용을 줄일 수 있다. 본 연구는 실제 서비스 [2]에도 적용하여 효과를 확인하였다.

2. 본론

2.1 시스템 아키텍처



(그림 1) (시스템 모듈)

제안된 시스템의 구성은 그림1과 같다.

*Caching Module: 클라이언트가 데이터 요청 시, 먼저 캐시 메모리를 확인하여 해당 데이터가 있는지 확인한다. 이 모듈은 2.2의 데이터 흐름을 관리한다.

*encrypt Module: 캐시에 저장된 데이터를 AES-256 알고리즘으로 암호화한다.

*Listener Module: 클라이언트의 캐시에 데이터가 저장되어 있고, 동시에 서버에서 데이터 변경이 발생할 경우, 캐시를 최신 데이터로 갱신하는 기능을 추가했다.

*Json Module: 특정 클라우드 서비스에서 기본 자료형 외에 자체 개발된 자료형을 지원하지 않는 문제를 해결한다.

2.2 데이터 흐름

*데이터 Get: 먼저 캐시 메모리를 확인하여 데이터가 존재하면, 서버 요청 없이 캐시에서 데이터를 불러온다. 데이터가 없거나 최신 데이터가 필요한 경우, 서버에 데이터를 요청한다. 서버에서 데이터가 도착하면 이를 클라이언트 캐시 메모리에 저장하고, 암호화 모듈을 이용해 암호화한다.

*데이터 Set: 데이터는 서버로 전달되고, 서버는 이를 수신하여 저장한 후 데이터 갱신 사실을 클라이언트에게 알린다. 클라이언트는 갱신된 데이터를 캐시 메모리에 저장한다.

2.3 암호화 및 보안

클라이언트의 암호화 모듈은 AES256 알고리즘을 사용하여 데이터를 암호화하고 복호화한다. 암호화된 데이터는

클라이언트 캐시 메모리에 저장된다.

2.4 데이터 요청 최소화

제안된 시스템은 클라이언트와 서버 간 데이터 요청을 최소화하도록 설계되었다. 데이터 유효성을 검사하고 서버에서 데이터 변경이 발생할 때 클라이언트에 이를 알림으로써 최신 데이터를 캐시에 반영하여, 서버 요청을 줄이고 전체 시스템 효율성을 높인다.

2.5 본 연구의 차별성

*클라이언트 측 캐싱을 통해 서버 요청 빈도를 줄여 서버 부하를 경감시키고, 클라우드 서비스의 운영 비용을 절감한다.

*AES-256 암호화 알고리즘을 사용하여 클라이언트 측 데이터의 보안성과 무결성을 보장하고, 클라이언트와 서버 간 데이터 흐름을 최적화하여 시스템 효율성을 향상시킨다.

*기존 연구와 같은 자체 개발 서버가 아닌 클라우드 백엔드 서비스와 연동된 시스템을 개발 및 운용하여 데이터 저장 및 접근 속도, 보안성, 비용 절감 효과를 종합적으로 평가하고 분석했다.

3. 구현 및 실험결과

3.1 서버 요청 빈도 감소:

표 1은 클라이언트 측 캐싱을 적용하지 않은 경우와 적용한 경우의 서버 요청 빈도를 측정한 결과이다. 클라이언트 측 캐싱을 적용한 결과, 서버 요청 빈도가 평균 66% 감소하였다.

(표 1) (캐싱 적용 전/후 요청 수 비교)

요청종류/캐싱적용여부	캐싱미적용	캐싱적용
Get 요청 수	14,000	2,800
Set요청 수	3,000	3,000
합계	17,000	5,800

3.2 응답 시간 감소

표 2는 클라이언트 측 캐싱 적용 여부에 따른 응답 시간을 보여준다. 캐싱을 적용했을 때 평균 응답 시간이 캐싱 미적용 대비 0.35% 수준으로 줄어들었다.

(표 2) (캐싱 적용 전/후 상세 시간 비교)

비교군	캐싱 적용	캐싱 미적용
cached check	0.0005ms	0.0005ms
Decryption	0.013ms	0.013ms
Deserialization	0.006ms	0.006ms
Firestore load	-	212.8ms
Total execution time	0.741182ms	213.5606ms

3.3 게임 진행 시간에 따른 캐싱 적중률 변화

게임 플레이를 분석한 결과, 시간이 지날수록 캐시 적중률이 점차 상승하는 경향을 보였다. 이는 게임이 진행될수록 자주 사용하는 데이터가 캐시에 저장되기 때문이다. 적중률이 증가하는 패턴은 자주 사용하는 데이터가 제한적임을 보여준다.

(표3). (게임 시간 별 캐시 적중률)

	10초	30초	1분	3분	10분	30분
적중률	98.69	98.74	99.41	99.45	99.99	99.99
	244%	386%	535%	345%	342%	994%

3.4 보안성 분석

일반적인 변수는 '메모리 서칭(Memory Searching)' 기법을 통해 쉽게 찾아내어 조작될 수 있다. GameGuardian과 같은 툴을 사용하면 이러한 변조가 가능하다. 그러나, 제안된 캐시 로더 시스템은 AES-256 알고리즘을 사용해 데이터를 암호화하며, 이 암호화된 키와 값을 사용해 변수를 관리하기 때문에 메모리 서칭을 통한 조작이 어렵다.

3.5 실제 상용화 결과

표 4는 게임[2]을 플레이스토어 및 앱스토어에 출시하고 데이터를 수집한 결과를 나타낸다. 일반적으로 대부분의 게임에서는 서버로부터 데이터를 가져오는 Get 요청이 Set 요청보다 더 자주 발생한다. 그러나, 이 게임에서는 Set 요청이 더 많이 발생한 것으로 나타났다. 이는 클라이언트 캐싱 기법이 효과적으로 작용해 클라이언트 측에서 캐시된 데이터를 잘 활용한 결과이다.

(표4). (게임[3] 클라우드 시스템 사용량 통계)

	1일	2일	3일	4일	5일	6일	7일
읽기	8,100	12,000	9,600	6,000	7,300	5,600	4,100
쓰기	10,000	13,000	9,900	8,700	11,000	6,900	6,000

참고 문헌

- [1]Google Firebase, <https://firebase.google.com/docs/firestore/billing-example?hl=ko>
- [2]CAOX, “모여라친구들 - 랜덤디펜스” <https://play.google.com/store/apps/details?id=com.CaoxGames.GatherFriends&hl=ko>

저학년 코딩교육을 위한 가상현실 기반 학습 방법 연구

홍서연¹, 박현아², 최지영³, 최미서⁴, 김장환⁵, 서채연⁶, 김영철^{7*}
^{1,4}홍익대학교 디자인컴퓨터전공학부,
^{2,3}홍익대학교 게임소프트웨어학과,
^{5,6,7*}홍익대학교 소프트웨어융합학과
 alicesyhong@gmail.com¹, pha2168@gmail.com², choejiyeong564@gamil.com³,
 miseo040415@gmail.com⁴, {lentoconstante⁵, chaeyun⁶, bob^{7*}}@hongik.ac.kr

Virtual Reality based Learning Approach for Coding Education in Lower Grades

Seo Yeon Hong, Hyeon-A Park, Ji Yeong Choe, Mi Seo Choi,
 Janghwan Kim, R. Young Chul Kim
 School of Design Convergence, Hongik University
 School of Games, Hongik University
 Dept. Software and Communications Engineering, Hongik University

요 약

최근, 2025년부터 코딩 교육이 초·중등학교에서 의무화됨에 따라, 효과적인 교육 방법에 대한 니즈가 높아진다. 기존의 텍스트 코딩의 어려움으로, 2차원 평면의 모니터를 통한 블록 코딩 방식이 초기 코딩 접근에 유용 했지만, 심화 학습과 흥미 유발에 한계가 있다. 이런 문제의 해결책으로, 가상현실 기술을 활용하여 3차원 공간에서 체험적 코딩 학습 방법을 제안한다. 게임 요소와 디지털 스토리텔링을 결합해 학습자에게 몰입감 있는 교육 환경을 제공하고, 창의적 문제 해결 능력과 컴퓨팅 사고력을 향상에 목적을 둔다. 제안하는 VR 기반 학습 방법의 프로토타입도 소개하며, 코딩교육의 접근성을 높이고 지속적인 학습을 유도할 가능성을 언급한다. 향후 연구로는 제안하는 학습 방법을 더욱 체계적으로 확장한다.

1. 서 론

이 논문은 메타버스 융합 SW 아카데미 (<https://metaverse.hongik.ac.kr>)의 교육 결과물이다. 2025년부터 시행되는 정부의 새로운 교육 정책은 초·중등 교육과정에서 코딩교육의 의무화를 명시한다. 이러한 변화로 인해 '정보' 과목의 수업 시간이 대폭 확대되어, 초등학교에서 34시간, 중학교에서 68시간으로 증가한다[1]. 이러한 교육 환경의 변화는 특히 저학년 학생들을 위한 효과적인 코딩 교육 방법이 요구된다. 기존의 2차원 블록 코딩 방식은 학습 초기 단계에서 유용하지만, 심화 학습과 지속적인 흥미 유발에는 한계를 나타낸다[2]. 이에 본 논문에서는 저학년 코딩교육을 위한 가상현실 기반 학습 방법을 제안한다. 가상현실 기반의 코딩교육은 학습자가 3차원 공간에서 직접 코드의 실행 결과를 체험하게 하여 추상적인 프로그래밍 개념을 시각적이고 체험적으로 이해하도록 돕는다[3]. 이러한 접근은 저학년 코딩교육에 흥미를 유도하기 위해 게이미피케이션 요소를 포함하여 학습 동기를 지속해서 유지할 수 있게 한다. 이에 따라 창의적 문제 해결 능력과 컴퓨팅 사고력을 높일 수 있기를 기대한다. 이를 통해, 저학년 학생들을 대상으로 한 가상현실 기반 코딩교육 프로

그램을 개선하여 초·중등 소프트웨어 교육에서 부족했던 개념에 대한 이해도를 높이고, 체험을 통한 학습자 중심의 프로그래밍 경험을 제공한다.

2장에서는 관련 연구로 기존의 저학년 코딩교육 방법과 기존 가상현실 기반 코딩교육에 대해 언급하고 3장에서는 제안하는 방법을 소개하고 그 과정에 대해 언급한다. 4장에서는 결론 및 향후 연구에 대해 언급한다.

2. 관련 연구

2.1 해외 코딩 기초 교육 사례



그림 1. 초등학생용 code.org 블록 코딩 예시
 Code.org는 전 세계 학생들이 컴퓨터 과학을 쉽게 배울 수

있도록 돕는 비영리 교육 플랫폼이다. 2013년에 설립된 Code.org는 주로 초등학교부터 고등학교까지의 학생들에게 컴퓨터 과학 교육을 제공한다. "모든 학생이 코딩을 배울 기회를 얻어야 한다"라는 목표를 가지고, 다양한 무료 온라인 학습 자료와 커리큘럼을 제공한다[4]. Code.org는 블록 기반 코딩, 텍스트 기반 프로그래밍, 웹 개발 등 다양한 학습 과정을 지원한다. 그림1은 코딩 메커니즘을 블록으로 나타내 학습하는 과정의 사례를 나타낸다.

2.2 가상현실 기반 코딩게임 사례

가상현실(VR) 기반의 블록 코딩을 학습할 수 있는 스팀(Steam) 플랫폼에서 제공하고 있는 'Koorring'은 2020년 제작된 메타 교육 콘텐츠이다[5]. Koorring은 Code.org와 다르게 캐릭터가 움직이는 시점을 1인칭으로 제작해 반복블록을 다양하게 활용할 수 있도록 했으며 VR을 통해 맵을 한 눈에 확인할 수 있어 사각지대없이 맵을 파악하기 쉽다는 장점이 있다. 그림 2는 Koorring 블록코딩 화면을 나타낸다. Koorring은 게임으로 제작되었지만 학습자의 위치를 고정시켜 흥미보단 학습에 집중한다.



그림 2. 블록코딩게임 Koorring 게임실행 화면

3. 저학년 코딩교육을 위한 가상현실 기반 학습 방법

본 논문은 가상현실 기술을 활용해 코딩 학습 환경을 실감형 콘텐츠를 제공한다. 3D 가상공간에서 블록을 통한 코딩, 게이미피케이션, 디지털 스토리텔링을 결합하여 학습자에게 체험적인 실감형 교육 콘텐츠를 제공한다.



그림 3. 마법진과 블록 스펬을 통한 체험적 코딩 학습 프로토타입
그림3은 본 논문에서 제안하는 가상현실 기반 실감형 블

록 코딩 프로토타입을 나타낸다. 학습자는 주어진 문제를 보고 블록을 마법 스펬, 즉 블록을 마법진에 배치한다. 각각의 코딩 블록은 한 줄의 코드로, 코딩 블록의 나열은 코드가 프로그램을 수행하는 메커니즘과 맵핑된다.

마법진이 완성되면 '루니'가 움직여 주어진 코딩블록의 기능을 순차적으로 수행하는 것은 일반적인 코딩에서 논리적으로 배치된 코드가 수행하는 것과 맵핑된다. 만약, 코딩 블록 입력 오류 시 뒤로가기 버튼을 통해 배치된 코딩블록을 수정할 수 있다. 블록은 이동블록과 반복블록으로 나뉜다. 이동블록은 전진, 왼쪽 회전, 오른쪽 회전, 올라가기, 아래로 내려가기 기능이 있다. 반복 블록은 2개의 블록사이 들어간 블록을 최대 4번 반복할 수 있다.

주어진 문제는 8개로 점진적으로 어려워지며, 각 문제에 맞는 퀘스트가 주어진다. 학습자는 튜토리얼을 통해 VR기기의 컨트롤러 사용법과 블록코딩으로 문제를 해결하는 방법을 학습한다. 마지막으로, 학습자는 결과를 확인하고 별점 위젯을 통해 피드백을 받는다.

4. 결론

본 논문은 VR 기술의 실감형 콘텐츠를 활용하여 코딩 메커니즘을 체험적으로 학습하는 새로운 방안을 제시한다. 제안하는 방법은 3D 가상공간에서 블록코딩과 게이미피케이션을 결합하여, 학생들에게 몰입감 있고 직관적인 학습 환경을 제공한다. 이를 통해 코딩 교육의 접근성을 높이고, 학습자의 지속적인 학습을 유도한다. 향후, 가상현실 기반 체험적 학습 방법을 확장하여 체계적인 코딩교육 콘텐츠에 관해 연구한다.

ACKNOWLEDGEMENT

본 연구는 2024년도 문화체육 관광부의재원으로 한국콘텐츠진흥원 (과제명: 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술 개발, 과제번호: RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초산산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임.

참고문헌

[1] 이주철 ,2025년부터 초등학교에서 '컴퓨터 언어' 코딩 배운다, 더리포트, 2022.08.22, <https://www.thereport.co.kr/news/articleView.html?idxno=23172>
 [2] 강혜영. (2022). 유아 코딩 교육에 대한 학부모 관심도 및 인식 조사 [석사학위논문, 공주교육대학교]. <http://www.riss.kr/link?id=T16141704>
 [3] 이지혜. (2018). VR시스템 환경 기술과 사용자 몰입감 요소 분석. 한국디자인문화학회지, 24(2), 585-596, 10.18208/ksdc.2018.24.2.585
 [4] Code.org. 2013. Learn Computer Science. Code.org. <https://code.org>
 [5] Vrani. (n.d.). Koorring. Vrani Inc. <http://vrani.co.kr/>

Video Swin Transformer를 이용한 올바른 운동 자세 추정

윤혜성, 백으뜸

호남대학교 AI 빅데이터 학과

e-mail : remisen326@naver.com, geodo100@gmail.com

Estimating correct exercise form using Video Swin Transformer

Hye-Seong Yoon, Eu-tteum Baek
Honam University, AI&Bigdata

요 약

잘못된 운동 자세는 신체의 불균형과 부상을 야기할 뿐만 아니라 운동의 효율성을 크게 떨어뜨려 운동자의 시간을 낭비시키고 원하는 결과를 얻기 어렵도록 만든다. 최근 몇 년간 컴퓨터비전 분야에서는 헬스케어 기술에 대한 관심이 집중되고 있다. 본 논문에서는 올바른 운동 자세 추정 문제에 Video Swin Transformer를 적용하고 전이학습과 파인튜닝을 진행하여 적용 가능성에 대해 검토하고자 한다. 실험 결과, 학습 횟수 25회만으로 Accuracy 0.85, Precision 0.60, Recall 0.58, F1 Score 0.59를 달성했다.

1. 서 론

자신이 올바른 자세로 운동하고 있는지 알 수 없어 불편함을 경험하고 있는 사람이 늘고 있다. 잘못된 운동 자세는 부상을 초래할 수 있을 뿐만 아니라 운동의 효율성을 크게 떨어뜨려 많은 시간을 투자해도 원하는 결과를 얻기 어려울 수 있다. 올바른 자세로 운동하고 잘못된 운동 자세를 올바르게 교정하는 것은 중요하다.

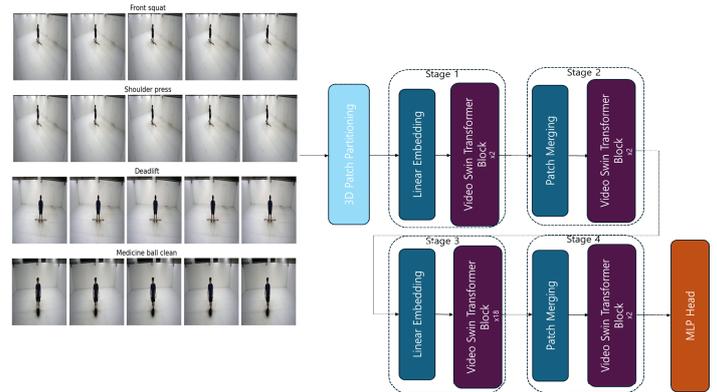
최근 몇 년간 컴퓨터비전 분야에서는 인간의 동작과 자세를 정확하게 분류하는 기술에 대한 연구가 활발하게 진행되고 있다. 이 기술은 헬스케어, 스포츠 분석, 가상 현실, 수화 번역, 인간과 로봇간의 상호작용에 사용되는 등 다양한 분야에서 활용이 가능하기에 그 중요성이 더욱 부각되고 있다.[1] 자세히 예를 들자면, 헬스케어 분야에서는 환자의 재활 치료 과정을 모니터링하고 평가하는 데 있어 필수적이고[2], 스포츠 분석에서는 선수들의 움직임을 분석하여 경기력을 향상시키고 부상을 예방하는 데 활용될 수 있다[3]. 가상현실과 증강현실에서는 사용자와의 상호작용을 자연스럽게 만들기 위해 정확한 동작 인식이 요구된다[4].

Video Swin Transformer는 Swin Transformer 모델을 확장하여 만들어진 모델이다. 인간 자세 추정은 2D 이미지나 비디오 프레임으로부터 특징을 추출한 뒤 이를 CNN 모델로 학습시키는 방법이 전통적으로 주로 사용되었다[5]. CNN은 영상 인식을 위한 필수적인 아키텍처였으나 대규모 데이터 세트에서 확장성과 성능 면에서 Video Swin Transformer가 CNN보다 더 우수함을 증명했다.[6] 뿐만 아니라 Video Swin Transformer는 비디오 시퀀스 내에서의 시간적 정보와 공간적 정보를 동시에 고려할 수

있으며 계층적 설계와 윈도우 기반 어텐션 메커니즘이 도입되어 고해상도 영상에서도 효율적인 처리와 정확한 분석이 가능하다.

본 연구에서는 Video Swin Transformer를 활용한 올바른 운동 자세 추정 방법을 제안한다. Video Swin Transformer의 자세 추정 가능성 검토를 위하여 사전훈련된 Video Swin Transformer를 전이학습하고 파인튜닝하여 자세 추정 문제에 적용한 실험 결과를 담아낸다.

2. Video Swin Transformer를 이용한 올바른 운동 자세 추정 방법



(그림 1) 제안하는 방법

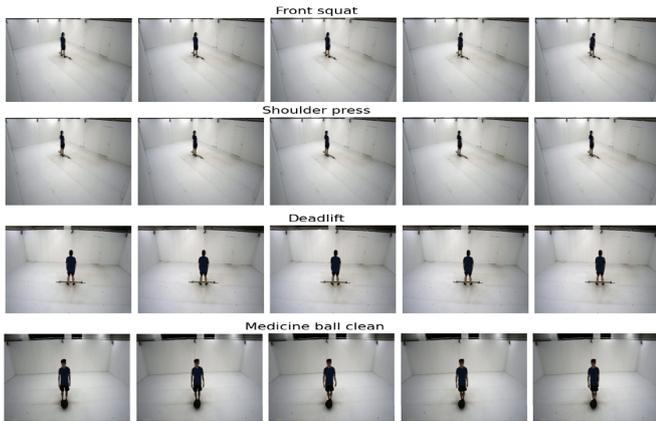
Video Swin Transformer에는 몇 가지 버전이 있는데, 본 절에서는 Swin-B를 이용한 올바른 운동 자세 추정 방법에 대해 소개하고 있다. Swin-B를 사용하여 하이퍼파라미터의 개수는 128개이다. 그림 1은 Video Swin

Transformer를 도식화한 것이다. 사용한 운동 자세 데이터에 대한 설명은 3. 실험 결과에서 다룬다.

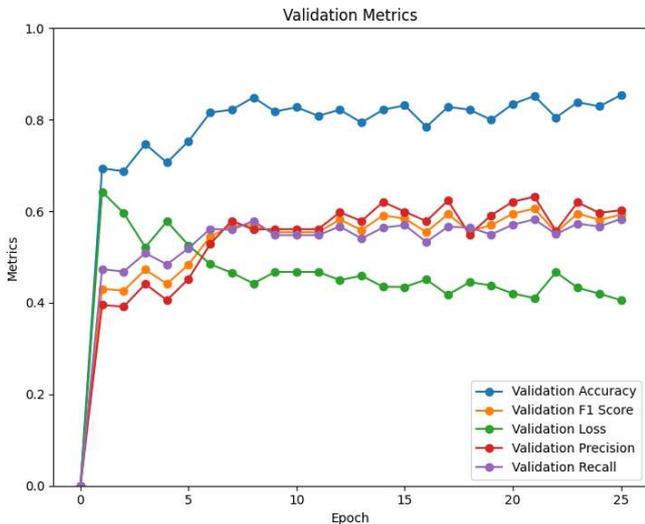
모델은 사전훈련된 videoswin_base_kinetics600_imagenet22k를 전이학습하여 파인튜닝하였다. 모델의 Input Shape는 32, 224, 224, 32이고 Window size는 8,7,7로 설정해주었다. 학습률은 1e-3, Weight decay 1e-4, AdamW 옵티마이저, 손실함수로는 Sparse Categorical Crossentropy를 사용하여 훈련하였다. 성능평가지표로는 Accuracy, F1 Score, Precision, Recall을 사용하였다.

3. 실험 결과

4.1 데이터셋 및 실험 결과 그래프



(그림 2) 포즈 데이터셋



(그림 3) 실험 결과 그래프

Model	Accuracy	Precision	Recall	F1 Score
Proposed Method	0.85	0.60	0.58	0.59

(표 1) 실험 결과 표

제안한 모델은 AIHUB에서 제공하는 크로스핏 동작 데이터 세트 중 일부를 사용하여 검증하였다. 크로스핏 동작 데이터 세트는 크로스핏 운동자 남녀를 다양한 각도에서 촬영한 데이터 세트이다. Front squat, Shoulder press, Deadlift, Medicine ball clean 4가지 운동 자세 동영상을 사용하였다. 사용한 동영상의 총 개수는 2983개이고 2655개는 훈련용, 328개는 검증용으로 분리시켰으며 각 비디오 클립을 224x224 해상도의 32 프레임으로 리샘플링하였다.

학습 횟수는 25회이며 Validation Accuracy 0.85, Precision 0.60, Recall 0.58, F1 Score 0.59를 달성했다. 낮은 학습 횟수임에도 불구하고 상당히 만족스러운 성능을 보였다. 낮은 학습 횟수를 고려했을 때, 학습 횟수를 늘릴 필요가 있는 것으로 보인다.

4. 결론

Video Swin Transformer는 적은 학습 횟수로도 좋은 정확도를 달성하며 운동 자세 추정 작업에 대한 좋은 잠재력을 입증했다. 향후 연구에서는 추가적인 데이터 세트 탐색, 학습 횟수 추가, 비교군 추가, 아키텍처 최적화를 통해 성능을 더욱 향상시키는 데 초점을 맞출 예정이다.

감사의 글

본 과제는 과학기술정보통신부의 출연금으로 수행한 AI융합대학 지원사업 연구결과입니다.

참고 문헌

- [1] GONG, Jia, et al. Diffpose: Toward more reliable 3d pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 13041-13051.
- [2] MASKELIŪNAS, Rytis, et al. BiomacVR: A virtual reality-based system for precise human posture and motion analysis in rehabilitation exercises using depth sensors. Electronics, 2023, 12.2: 339.
- [3] LUDWIG, Katja, et al. Self-supervised learning for human pose estimation in sports. In: 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2021. p. 1-6.K. Elissa, "Title of paper if known," unpublished.
- [4] KONDORI, Farid Abedan, et al. Direct hand pose estimation for immersive gestural interaction. Pattern Recognition Letters, 2015, 66: 91-99.
- [5] MUNEA, Tewodros Legesse, et al. The progress of human pose estimation: A survey and taxonomy of models applied in 2D human pose estimation. IEEE Access, 2020, 8: 133330-133348.
- [6] LIU, Ze, et al. Video swin transformer. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022. p. 3202-3211.

로컬 어텐션 기반의 밀집 해충 카운팅 모델

조정민, 김경태, 최원준, 손창환(교신저자)*
 군산대학교 소프트웨어학과, 군산대학교 소프트웨어학과*
 e-mail : 5281cjm@gmail.com, 3dlife4620@naver.com, cwj000823@gmail.com,
 cson@kunsan.ac.kr*

Dense Pest Counting Model Based on Local Attention

Jeong-Min Cho, Kyoung-Tae Kim, Won-Jun Choi and Chang-Hwan Son*
 Department of Software Science & Engineering, Kunsan National University

요약

트랩에서 촬영된 해충 이미지는 해충이 밀집되거나 군집을 형성하기 때문에 폐색 문제 및 다양한 포즈 변화를 야기한다. 이에 대해 본 연구는 스택 백본의 첫 번째에서 예측된 해충의 중심점 등의 정보를 활용한 로컬 어텐션 설계를 통해 특징 구별력을 강화하는 방법을 제안한다. 실험을 통해 폐색 및 포즈 변화 문제를 효과적으로 극복하였고 제안된 모델은 로컬 그룹화 및 차별적 특징 어텐션을 통해 특징 구별력을 향상하였고, 폐색 및 포즈 변화 문제를 극복하는 데 효과적인 것으로 입증되었다.

1. 서론

해충은 작물의 품질 저하와 경제적 손실을 초래하며, UN 식량 농업 기구에 따르면 매년 작물 생산량의 약 40%가 해충으로 손실된다[1]. 이를 해결하기 위해, 최근 디지털 트랩을 노지에 설치하여 해충 수를 파악하여 조기 방제를 적용할 수 있는 무인예찰 시스템이 도입되고 있다. 하지만, 트랩에 포획된 해충이 고밀도로 분포하여 해충 간에 서로 가려지는 폐색 문제, 포즈의 다양화, 해충 외관의 유사성 문제가 딥러닝 모델을 개발하는 데 걸림돌로 작용하고 있다. 따라서 본 연구에서는 밀집된 해충 카운팅 문제를 해결할 수 있는 새로운 해충 카운팅 모델을 제안하고자 한다.

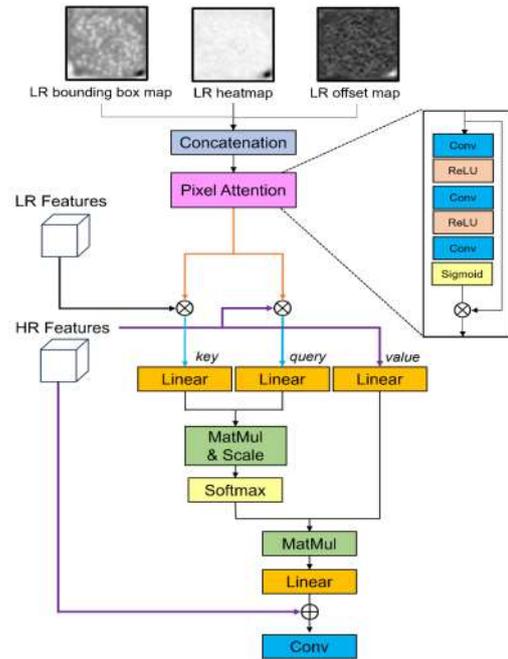
2. 제안한 로컬 어텐션 기반의 해충 카운팅 모델

기존 멀티스케일 센터넷[2]은 LR(Low-Resolution) 및 HR(High-Resolution) Hourglass라 불리는 서로 다른 스케일의 백본과 전역 어텐션으로 HR 예측치가 다운샘플링에 따른 이산적인 오차를 줄여 해충의 폐색 문제에 더 강인하게 대처할 수 있게 하였는데, 이는 객체와 배경에 대한 그룹화가 되어있지 않아서 해충 객체의 폐색이 심한 경우에는 해충 객체를 정확하게 분리해 내는 데에 한계가 있다. 따라서 본 연구에서는 객체 정보를 그룹화하여 전경과 배경을 분리하고 HR 특징의 구별력을 강화하는 것을 목적으로 하는 로컬 그룹화 및 스케일 유도 어텐션을 제안한다.

2.1 학습 가능한 공간 어텐션 맵 변환

제안된 로컬 어텐션 모듈은 그림 1의 상단에서 알 수 있듯이 예측 맵을 입력으로 필요로 한다. 특히, 히트맵은 해충의 중심점을 포함하고 있기에 로컬 그룹화를 위한 초기 시작점으로 활용될 수 있다. 하지만, 히트맵 자체는 예측치로 오차를 포함할 수 있으며 최소한 분포를 가지기 때문에 히트맵을 바로 사용하는 것은 바람직하지 않다. 따라서 제안한 로컬 어텐션 모듈에서는 히트맵을 포함한 예측맵을 학습 가능한 공간 어텐션 맵으로 변환하고자

한다. 이를 위해, 픽셀 어텐션 모듈[3](Pixel Attention Module : PAM)을 적용하고자 한다.



$$F = \text{Concat}(\hat{Y}^{LR}, \hat{S}^{LR}, \hat{O}^{LR}) \quad (1)$$

$$W = f^s \circ f^c \circ f^a \circ f^c \circ f^a \circ f^c(F) \quad (2)$$

$$M = F \otimes W \quad (3)$$

수식 1의 \hat{Y}^{LR} , \hat{S}^{LR} , \hat{O}^{LR} 는 각각 예측된 히트맵, 바운딩 박스 맵, 오프셋 맵을 나타낸다. 그리고 수식 2의 f^s , f^c , f^a 는 각각 시그모이드, Conv, ReLU 계층을 의미한다. 수식 1은 이 세 가지 예측 맵을 채널 축을 따라 쌓아서 PAM의 입력 텐서 F 를 생성한

다. 수식 2는 입력 텐서 F 에서 어떤 영역을 더 강조하고 억제할지를 제어하기 위한 2차원 가중치 맵 W 를 생성한다. PAM의 최종 출력은 입력 F 와 가중치 맵 W 의 요소별 곱셈인 수식 3의 결과이다. 이 수식은 로컬 그룹화의 초기 솔루션인 예측 맵을 학습 가능한 공간 어텐션 맵으로 변환한 것을 의미한다. 이는 로컬 특징 그룹화를 보다 유연하고 정확한 방향으로 유도할 수 있다.

2.2 로컬 그룹화와 스케일 유도 어텐션

제안한 모델에서는 객체 특징을 추출하기 위해, 이미지 필터링 개념을 사용하여 학습 가능한 공간 어텐션 맵인 M 을 입력 LR 및 HR 특징에 곱하여 객체 영역을 추출하고 객체 특징에서 의미 정보를 추출하기 위해 트랜스포머에서 소개된 자가 어텐션(Self Attention) 모델[4]을 이용해 로컬 어텐션을 모델링하고자 한다.

$$Q_F = Y(\Phi_{HR}^l) \otimes f^{d(l)}(M) \quad (4)$$

$$K_F = Y(\Phi_{LR}^l) \otimes f^{d(l)}(M) \quad (5)$$

$$V = Y(\Phi_{HR}^l) \quad (6)$$

Φ_{HR}^l 과 Φ_{LR}^l 은 각각 HR과 LR hourglass의 l 번째 잔차 블록에서 추출한 HR과 LR 특징을 나타낸다. 그리고 Y 는 임베딩 계층이며 $f^{d(l)}$ 는 다운샘플링을 의미한다. 수식 4와 5에서 학습 가능한 공간 어텐션 맵인 M 이 각각 HR 특징과 LR 특징에 원소별로 곱해져 이미지 필터처럼 작용하여 객체 특징을 추출하고 배경 특징을 제거하는 역할을 한다. 그리고 트랜스포머에서 사용된 자가 어텐션과는 달리, 수식 4-6에서 쿼리, 키, 값이 다른 형태를 취한다. 필터링 된 HR과 LR 객체 특징은 Q_F 와 K_F 로 사용되고, HR 특징은 V 로 사용된 것을 볼 수 있다.

$$Atten(Q_F, K_F, V) = \text{softmax}\left(-\frac{Q_F K_F^T}{\sqrt{d_k}}\right) V \quad (7)$$

여기서 $Q_F K_F^T$ 은 유사도 행렬을 의미하고, 필터링된 HR 특징 Q_F 와 필터링된 LR 특징 K_F 가 얼마나 유사한지 그 상관성을 측정하여 저장하고 있다. 이 유사도 행렬은 수식 7에서와 같이 HR 특징 V 의 선형 결합을 위해 가중치 값으로 사용된다.

$$\Phi_{HR}^{(l+1)} = f^c(\Phi_{HR}^l + Atten(Q_F, K_F, V)) \quad (8)$$

로컬 어텐션이 적용된 결과는 수식 8와 같이 스킵 연결을 통해 입력 HR 특징과 원소별 덧셈을 통하여 푸진되고 최종 합성곱 계층을 통과한다. HR 특징 $\Phi_{HR}^{(l+1)}$ 을 다음 HR Hourglass 계층으로 전달하는 과정을 반복하여 특징 구별력을 향상시키고 궁극적으로 밀집된 해충 카운팅의 성능 개선으로 이어질 수 있다.

3. 실험 결과

제안한 로컬 어텐션 모델을 평가하기 위해, 기존모델들과 비교하였다. 정량적 지표를 위해, 평균 절대 오차(Mean Absolute Error : MAE), 평균 제곱근 오차(RMSE, Root Mean Squared Error : RMSE), 평균 정밀도(Average Precision : AP)를 계산하였다. MAE와 RMSE 지표는 해충 카운팅의 정확도를 반영할 수 있고 AP 지표는 바운딩 박스 검출의 정확도를 측정할 수 있다.

표 1은 해충 카운팅 모델에 대한 정확도 결과를 보여준다. 표 1에서 새로운 해충 카운팅 모델 개발이 폐색 문제, 포즈 가변성,

외관 유사도 문제 등에 대해 기존모델들에 비해서 더욱 효과적으로 작용하여 성능이 크게 개선된 것을 확인할 수 있다.

(표 1) 해충 카운팅 모델의 정량적 평가

Counting Models	MAE(↓)	RMSE(↓)	AP(↑)
RetinaNet[5]	39.589	52.454	0.595
Faster RCNN[6]	29.397	42.553	0.521
RepPoints[7]	28.404	43.095	0.678
YOLOX[8]	11.612	14.717	0.676
YOLOv7[9]	4.843	16.573	0.609
CenterNet[10]	6.015	11.031	0.653
Multiscale CenterNet[2]	3.334	9.719	0.693
Proposed Model	2.538	6.666	0.920

ACKNOWLEDGMENT

이 연구는 농촌진흥청 연구사업(과제번호: PJ01630302)의 지원을 받아 수행된 연구임

참고 문헌

- [1] W. Zhang, H. Huang, Y. Sun, and X. Wu, "AgriPest-YOLO: A rapid light-trap agricultural pest detection method based on deep learning," *Frontiers in Plant Science*, vol. 13, Dec. 2022.
- [2] J.-H. Lee and C.-H. Son, "Trap-based pest counting: Multiscale and deformable attention CenterNet integrating internal LR and HR joint feature learning," *Remote Sensing*, vol. 15, no. 15, 3810, July 2023.
- [3] H. Zhao, X. Kong, J. He, Y. Qiao, and C. Dong, "Efficient image super-resolution using pixel attention," *arXiv:2010.01073v1 [eess.IV]*, Oct. 2020
- [4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems*, Long Beach, USA, Dec. 2017.
- [5] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE International Conference on Computer Vision*, pp. 2999-3007, Oct. 2017
- [6] .S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, June 2017.
- [7] Z. Yang, S. Liu, H. Hu, L. Wang, and S. Lin, "RepPoints: Point set representation for object detection," in *Proc. IEEE International Conference on Computer Vision*, pp. 9657-9664, Aug. 2019.
- [8] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO Series in 2021," *arXiv:2107.08430 [cs.CV]*, Jul. 2021.
- [9] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv:2207.02696 [cs.CV]*, Jul. 2022.
- [10] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," *arXiv:1904.07850v2 [cs.CV]*, Apr. 2019.

요양시설 내 노약자 낙상 감지 모니터링 시스템

박준렬¹, 박민현¹, 이민혁¹, 전민서¹, 양진모², 김영철^{3*}

¹홍익대학교 소프트웨어융합학과 학사과정

²홍익대학교 소프트웨어융합학과 석사과정

^{3*}홍익대학교 소프트웨어융합학과 교수

e-mail: {qkrwnsfuf123, minhyun102947}@gmail.com,

{alsgur2305, ms000320}@naver.com, yjmd2222@g.hongik.ac.kr, bob@hongik.ac.kr

Fall Detection Monitoring System for Elderly Patients in Nursing Home Facility

Junryeol Park¹, Minhyun Park¹, Minhyuk Lee¹, Minseo Jeon¹, Jinmo Yang²,
and R. Young Chul Kim^{3*}

Dept. of Software and Communications Engineering, Hongik University

요약

고령화 시대에 가정보다는 사회 및 정부가 많은 노약자를 돌봐야 하는 이슈가 중요하게 되었다. 그래서 요양시설의 역할이 강조되고 있는 현실이다. 이런 시설에서의 노약자 낙상 사고 또한 빈번히 일어나고 있다. 이를 해결하기 위해, You Only Look Once v8n-pose (YOLOv8n-pose) 모델 기반 노약자 낙상 감지 모니터링 시스템을 개발 중이며 이를 언급한다. 구체적으로, 실시간으로 고령자의 신체 주요 부위와 동작을 분석하여 낙상 여부를 감지하고, 위급 상황 발생 시 관리자가 즉시 대응할 수 있도록 알림을 전송하는 방법을 사용하고자 한다. 이를 통해 본 연구는 요양시설에서의 사고 예방 및 대응 능력을 향상시키는 것에 기여할 것을 기대한다.

1. 서론

본 논문은 2024년 2학기 홍익대학교 소프트웨어융합학과 종합설계의 프로젝트 결과물이다.

한국 사회는 급격한 고령화로 인해 70대 이상의 인구가 최초로 20대 인구를 초과하는 인구 구조 변화를 겪고 있다[1]. 이에 따라 요양시설 내 고령자들의 안전 문제는 점점 더 중요한 사회적 과제로 대두되고 있다. 특히, 요양시설에 보호 관찰자가 부재한 시간대에 발생할 수 있는 낙상 사고나 응급 상황에 대한 대비가 필수적이다. 따라서 본 연구는 YOLOv8n-pose 모델을 활용하여, 요양시설 내에서 고령자의 낙상 사고를 실시간으로 감지하고 신속하게 대응할 수 있는 시스템을 개발하는 것을 목표로 한다. 이를 통해 요양시설의 안전성을 높이고, 고령자들의 삶의 질을 향상하는 것에 기여할 수 있다. 2장에서는 기존의 낙상 관련 연구를 소개한다. 3장에서는 요양시설 내 노약자 낙상 감지 모니터링 시스템을 언급한다. 4장에서는 결론 및 향후 연구를 언급한다.

2. 관련 연구

낙상은 고령자와 거동이 불편한 환자들에게 심각한 위협 요소로, 이를 방지하기 위한 다양한 기술들이 연구되어 왔다. 특히, 영상처리 기반의 낙상 예방 모니터링 시스템은 최근 들어 많은 주목을 받고 있다.

2.1 영상처리 기반 환자 낙상 예방 모니터링 시스템

영상처리 기반 환자 낙상 예방 모니터링 시스템은 병원에서 거동이 불편한 환자의 낙상을 사전에 방지하기 위해, 카메라를 이용한 영상처리 기반의 낙상 예방 모니터링 시스템을 제안한다[2]. 이 시스템은 카메라와 각종 센서를 결합하여, 실시간으로 환자의 움직임을 감지하고, 이상 상황 발생 시 스마트폰 또는 PC를 통해 알림을 보낸다. 이 시스템은 관리자의 주의 없이도 낙상을 감지할 수 있지만, 고정된 침대 구도를 사용하기 때문에 다양한 병원 시설 직접적인 적용이 어려운 한계가 있다.

2.2 YOLOv5 기반 낙상사고 인식 연구

YOLOv5 기반 낙상사고 인식 연구는 YOLOv5 모델을 이용해 독거 노인의 낙상 사고를 인식하는 방법을 제안하였다[3]. 이 연구에서는 실제 낙상 상황을 재현한 이미지를 수집하고 라벨링하여 데이터셋을 구축한 후, YOLOv5 모델을 학습시켰다. 성능을 평가해본 결과 이 모델이 대부분의 낙상 사고를 정확하게 인식했다. 그러나 이 모델은 탐지할 객체에 바운딩 박스를 추론하여 객체를 식별하므로, 복잡한 인간의 자세 변화나 특정 행동 패턴을 인식하기에는 한계가 있다.

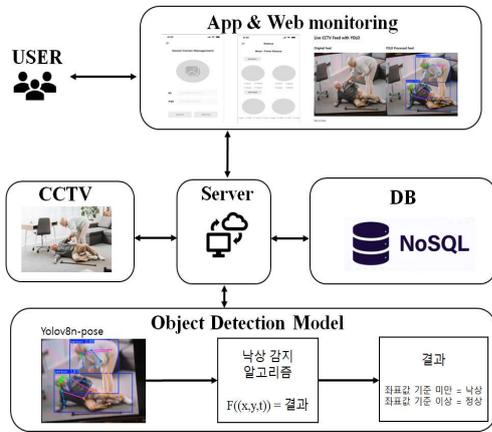
3. 본 론



(그림 1) YOLOv5와 YOLOv8n-pose 감지 결과 비교

본 연구는 YOLOv8n-pose 모델을 활용하여 사람의 동작을 정밀하게 인식한다. 이는 사람의 바운딩 박스 뿐만 아니라 주요 관절을 나타내는 좌표인 키포인트를 통한 자세 추정으로 진행된다[4]. 따라서 이를 활용한 요양시설 내 낙상 감지 모니터링 시스템은 사람의 뼈대 정보를 통해 낙상 과정을 정밀하게 감지할 수 있다.

그림 2는 이 모니터링 시스템에 대한 구상도이다.



(그림 2) 요양시설 내 낙상 감지 모니터링 시스템 구상도

이 시스템의 흐름은 아래와 같이 요약될 수 있다.

3.1 영상 수집 과정

영상 수집 과정은 요양시설 내 설치된 CCTV를 통해 고령자의 일상 활동을 실시간으로 촬영하고 데이터를 수집하여 서버로 전송한다.

3.2 낙상 감지 알고리즘

서버에 전송된 영상은 YOLOv8n-pose 모델을 통해 분석된다. 이 모델은 그림 1의 우측 사진과 같이 사람의 키포인트를 실시간으로 추론하여 분석하고 특정 임계값을 기준으로 키포인트가 감지될 경우 낙상 사고로 판단한다. 세부적인 판단 기준은 아래와 같다.

1) 급격한 위치 변화: 높이를 나타내는 y 좌표의 변화는 낙상의 과정을 표현할 것으로 간주한다. 이전 프레임과 현재 프레임에서의 y 좌표의 변화를 비교하여 그 값이 직전 프레임의 y값보다 33 % 작은 것을 초기 조건으로 한다.

2) 지속성: y 좌표의 급격한 변화 후 최소 5~10 초 동안 해당 위치에서 움직임이 없어야 한다. 이는 낙상 후 일어

날 수 없는 상황을 가정하여, 사람이 낙상으로부터 회복할 수 없는 것을 표현한다.

3) 탐지 불가능 상태의 지속: y 좌표의 급격한 변화 후 카메라 감지 범위를 벗어나는 상태가 최소 5~10 초 동안 지속될 경우를 확인한다. 이는 낙하 중 사람이 카메라의 시야에서 벗어나는 상황을 표현한다.

이러한 세부적인 분석 접근 방식은 기존 모델들이 제공할 수 없는 정밀한 데이터 처리능력을 보여준다.

3.3 응급상황 알림 과정

낙상이 감지되면, 시스템은 즉시 관리자에게 실시간 알림을 보낸다. 이 알림은 관리자의 스마트폰 또는 웹 기반 플랫폼을 통해 전달되며, 관리자는 즉각적으로 상황을 파악하고 적절한 조치를 할 수 있다. 신속한 대응은 고령자의 안전을 보장하고 사고로 인한 위험을 최소화한다.

4. 결 론

본 논문은 사람의 자세와 동작을 정교하게 인식할 수 있는 YOLOv8n-pose 모델을 활용하여 요양시설 내 고령자의 낙상 사고를 감지하는 모니터링 시스템의 개발 과정을 소개한다. YOLOv8n-pose 모델은 실시간 추론이 가능하고 사람의 구체적인 모습을 키포인트로 나타내기 때문에 요양시설 내 실시간으로 발생할 수 있는 낙상 사고를 보다 신속하고 정확하게 감지할 수 있다. 따라서 요양시설의 안전 관리를 한층 더 강화시킬 것을 기대할 수 있다. 현재 시스템은 아직 개발 단계에 있으며, 다양한 환경에서의 성능을 더욱 철저한 검증과, 외부 환경에서도 안정적으로 작동할 수 있도록 개선이 필요하다. 향후 연구에서는 이 시스템을 완성하여 성능을 확인할 예정이다. 궁극적으로 요양시설뿐만 아니라 어린이집, 호스피스와 같은 시설에서도 확대하여 공공의 안전을 강화하고자 한다.

Acknowledgment

본 연구는 2024년도 문화체육관광부의 재원으로 한국콘텐츠진흥원 (과제명: 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술 개발, 과제번호: RS-2023-00227917, 기여율: 50 %) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005, 기여율: 50 %)의 지원을 받아 수행된 연구임.

참고 문헌

[1] 통계청. “고령화지수”, 『장래인구추계, 2024년 7월 16일.
 [2] 김만우. “영상처리 기반 환자 낙상예방 모니터링 시스템”, 『학위논문(석사)』 경남대학교 대학원, 첨단공학과, 2017.
 [3] 정수연, 강민지, 이병영. “독거노인을 위한 YOLOv5 기반 낙상사고 인식 연구”, 한국콘텐츠학회논문지, Vol. 23 No. 11, PP. 83-89, 2023
 [4] Ultralytics, “Ultralytics YOLO Docs,” Ultralytics, 2024 (Updated), <https://docs.ultralytics.com/tasks/pose/>.

청각장애 아동을 위한 수화 교육용 온라인 프로그램 개발*

강성연, 김서영, 정현정, 이은서, 송지영
한남대학교 컴퓨터공학과
20210498@gm.hannam.ac.kr, jysong@hnu.kr

Developing an Online Program for Sign Language Education for Hearing-Impaired Children

Seongyeon Kang, Seoyoung Kim,
Hyeonjeong Jung, Eunseo Lee, Jiyoung Song
Dept. of Computer Engineering, Hannam University

요약

수어 교육기관의 부족은 오랫동안 청각장애 아동들과 그 가족들이 직면해 온 문제이다. 현재 국내에는 국·공·사립 특수학교가 194개교 있으나 이중 청각장애 특수학교는 전국에 13곳뿐이며, 이들 또한 대부분은 수도권에 존재한다. 청각장애 아동들은 수어 통역 등의 교육 편의를 지원받지 못하고, 수어를 배우고 싶어도 그 기회를 쉽게 얻지 못하고 있다. 본 논문에서는 위 문제점을 해결하기 위해 청각장애 아동을 위한 수화 교육용 온라인 프로그램 개발을 제안한다. 해당 프로그램은 아동의 발달과 학습 속도에 맞춰 AI 아바타를 활용해 양방향 번역이 가능한 수화 학습 기능을 제공한다. 이를 통해 청각장애 아동의 교육 접근성과 학습 흥미도를 높이고, 다양한 배움의 기회를 제공할 수 있다. 나아가, 청각장애인이 겪는 소통의 어려움 또한 해소하게 될 것으로 기대된다.

1. 서론

현재 국내 특수학교 194개교 중 청각장애 특수학교는 13곳뿐이며, 서울 4교, 경기/인천 3교 등 과반이 수도권에 집중되어 있다. 이에 청각장애 아동 2,907명 중 특수교육 지원센터에 다니는 아동은 572명으로, 나머지는 일반 학교에 재학 중인 것으로 나타났다[1]. 이는 청각장애 아동을 위한 전문 교육기관이 여전히 부족한 실정임을 보여준다.

수어를 사용하지 않는 일반 학교에서 청각장애 아동은 교사의 입 모양에만 의존해야 하기에 공부하기 쉽지 않다[2]. 그러나, 특수학교에서도 수어를 쓰는 교사에게 체계적인 교육을 받는다는 보장이 없다. 현재 특수교사 양성체제는 장애 영역별 교사 자격제도가 아닌 특수교육의 보편성에 기반한다[3]. 그 결과, 청각장애 특수교육기관 내 수어 통역 자격을 가진 교사의 비율이 고작 6%에 불과한 것으로 나타났다. 청각장애 아동은 교육 편의를 충분히 제공받지 못하고, 수어 학습의 기회를 쉽게 얻지 못하고 있다.

이에 본 연구에서는 청각장애 아동을 위한 수화 교육용 온라인 프로그램 개발 아이디어를 제안한다. 수화 교육용 온라인 프로그램으로 청각장애 아동에게 시간과 장소의 제약 없이 체계적인 수어 학습 기회를 제공하고자 한다.

2. 관련 연구

수화 교육과 관련하여 기존에 진행된 여러 연구와의 비교 분석을 통해 본 연구에서 제안하는 수화 교육용 온라인 프로그램의 우수성을 확인하고자 한다.

기존 연구에서는 수화 교육을 다양한 방식으로 시도해 왔으나, 몇 가지 한계점이 존재한다. 3D 수화 교육 앱[4]은 AI 아바타를 교육 요소로 활용하지만, 학습 흥미 유발 요소와 사용자 동작 인식 기능이 부족하다. 동화로 배우는 수화 게임(웹)[5]은 동영상상을 활용한 학습과 동화 및 퍼즐을 통해 아동의 학습 흥미를 유발하나, 동작 인식 기능은 존재하지 않아 실시간 학습 확인이 어렵다. 수화 교육 앱[6]은 퀴즈를 통한 학습 흥미 유발과 사용자 동작 인식 기능을 제공하지만, 단계별 학습 커리큘럼은 제공하지 않는다. 지문자 학습 게임(웹)[7]은 AI 아바타를 통한 학습과 게임을 활용하여 흥미를 유발하지만, 해당 연구 역시 커리큘럼을 제공하지 않는다. 마지막으로, 수어 교육 서비스 플랫폼(웹)[8]은 동영상 학습과 사용자 동작 인식 기능, 단계별 학습 커리큘럼까지 여러 기능을 제공하지만, 학습 활용 요소가 동영상 자료에 한정된다는 한계가 있다.

이에 비해, 수화 교육용 온라인 프로그램은 웹과 앱 환경 모두에서 AI 아바타, 동영상, 사진 등 다양한 요소를 활용한 수화 교육을 제공한다. 게임을 통해 아동의 학습 흥미를 유발하고, 사용자 동작 인식 기능으로 실시간 학습 확인이 가능하게 한다. 또한, 단계별 학습 커리큘럼을 제공하여 체계적인 수화 학습이 가능하도록 돕는다. 이처럼, 본 프로그램은 기존의 여러 연구보다 더 우수한 결과를 나타낼 것으로 기대된다.

3. 수화 교육용 온라인 프로그램 개발

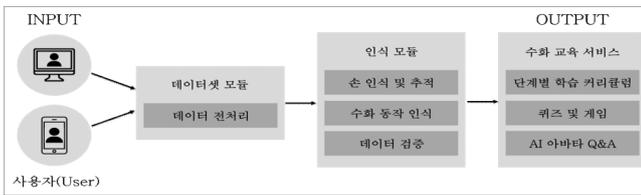
3-1. 수화 교육 서비스 기능 소개

* 이 논문은 2024년도 한남대학교 지방대학활성화 사업의 지원을 받아 수행된 연구임

본 연구에서 제안하는 수화 교육용 온라인 프로그램은 다음의 주요 기능들을 제공한다.

첫째, 단계별 학습 커리큘럼을 제공한다. 청각장애 아동의 발달과 학습 속도에 맞춰 단어부터 문장까지 체계적인 수화 학습이 가능하도록 돕는다. 둘째, 퀴즈 및 게임 기능을 제공한다. 아동에게 학습에 관한 흥미 요소를 제공하며, 아동이 학습 이후의 성취도를 직접 확인하고 수화를 숙지할 수 있게 한다. 셋째, AI 아바타 Q&A 기능을 제공한다. 아동이 학습한 수화를 양방향 번역이 가능한 AI 아바타와 소통하며 실상황에서의 수화 대화 능력을 기를 수 있도록 돕는다. 해당 기능들은 사용자의 수화 실력 향상과 수화에 대한 흥미 및 이해도 제고에 기여하게 될 것이다.

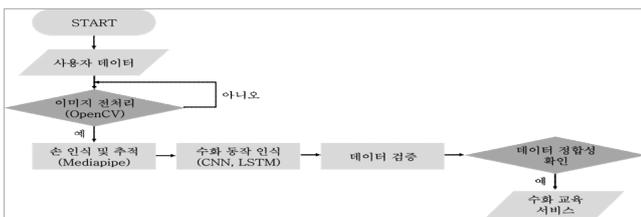
3-2. 프레임워크 구성



(그림 1) 수화 교육용 온라인 프로그램 프레임워크

그림 1은 본 연구의 프레임워크를 나타낸 것이다. 웹과 애플리케이션 환경에서 프로그램을 실행한 사용자로부터 손동작과 같은 사용자 데이터값을 입력(INPUT)받는다. 입력된 데이터값은 데이터셋 모듈로 이동하여 데이터 전처리 과정을 거친다. 전처리된 데이터는 인식 모듈에서 손의 좌표값과 관절 정보가 인식 및 추적된다. 이후, 전처리된 이미지를 기반으로 수화 동작을 인식하여 정확성을 확인하는 데이터 검증 과정을 거친다. 이 모듈들을 거친 데이터는 최종적으로 여러 기능을 가진 수화 교육 서비스로 제공된다는 출력(OUTPUT)값을 얻는다.

3-3. 프레임워크 구현 알고리즘



(그림 2) 프로그램 알고리즘 순서도

그림 2는 본 연구의 프레임워크 구현에 활용할 알고리즘의 순서도이다. 손동작과 같은 사용자 데이터를 입력받으면 OpenCV 오픈소스 라이브러리를 사용해 이미지 크기 및 색상 조정과 같은 전처리 과정을 거친다. 전처리 과정이 잘 이루어졌다면, MediaPipe 오픈소스 프레임워크를 사용해 전처리된 이미지를 분석하여 손의 좌표값 및 관절 정보를 추출한다. 추출된 데이터는 CNN과 LSTM을 활용한 딥러닝 과정을 통해 입력된 수화 동작의 인식 결과를 도출한다. 도출된 결과는 실제 수화 동작 정보가 저장된 모델과 데이터 정확성을 확인한다. 모순이 없다면, 이를 바탕으로 활용한 수화 교육 서비스 모듈이 제공되도록 하는 출력값을 얻는 순서로 알고리즘이 진행된다.

3-4. 개발 예상 결과물



(그림 3) 프로그램 기능별 개발 예상 결과물

그림 3은 수화 교육용 온라인 프로그램의 주요 기능별 개발 예상 결과물을 보여준다. 프로그램은 웹과 앱 환경 모두 연동되며, 직관적인 인터페이스로 사용자 편의를 제공한다. 단계별 학습 커리큘럼 기능은 사용자의 학습 속도에 맞춰 내용을 체계적으로 제공하며, 수화 게임 기능은 아동의 학습 흥미와 교육 접근성을 높인다. 마지막으로, AI 아바타 Q&A 기능은 사용자가 수화로 소통할 수 있는 환경을 제공하여 실전에서의 수화 대화 능력을 강화한다.

이와 같이 프로그램의 각 기능은 청각장애 아동의 수화 학습을 효과적으로 지원하게 될 것이다.

4. 결론

본 연구에서는 수화 교육용 온라인 프로그램의 개발 아이디어를 제안하였다. 수화 교육용 온라인 프로그램을 통해 청각장애 아동의 교육 접근성을 높이고, 다양한 학습 기회를 제공하고자 한다. 나아가, 본 프로그램이 청각장애 아동의 사회적 참여와 자립을 촉진하고 사회적 통합과 평등한 교육 환경 조성에 기여할 것으로 기대된다. 앞으로 본 프로그램의 실효성을 검증하고, 지속적인 개선을 통해 더욱 많은 청각장애 아동과 그 가족들에게 도움이 되는 방향으로 발전해 나가야 할 것이다.

참고 문헌

- [1] 교육부. (2023). 2023 특수교육 통계.
- [2] 박유빈. “전국 특수학교 192개교 중 농학교는 14곳 불과 [심층기획-말뿐인 공용어…실 곳 없는 한국수어(手語)]” *세계일보*. 2023-05-29.
- [3] 이재상. “정부, 무너진 농교육과 농학생 교육권 바로 세워 달라” *미디어생활*. 2023-09-21.
- [4] 정영기, “3D 수화교육 스마트폰 앱콘텐츠 개발”, *한국스마트미디어학회 스마트미디어저널*, 1(3), 2012.09, 8-14.
- [5] 김보란, 한준탁, “청각/언어 장애 아동을 위한 효율적인 수화 게임 구현”, *한국컴퓨터교육학회 2010년도 동계학술발표논문지*, 14(1), 2010.02, 217-222.
- [6] 김신영, 엄서정, 유선영, 김수정, 이경미, “미디어파이프와 장단기기억을 이용한 수화 동작인식 앱 개발”, *디지털콘텐츠학회논문지*, 24(1), 2023.01, 111-119.
- [7] 오영준, 정기철, “아바타를 내장한 한글 지문자 학습 게임 개발에 관한 연구”, *한국게임학회 논문지*, 9(4), 2009.08, 67-80.
- [8] 김미서, 유문희, 장지현, 최정윤, 나고은, 김미숙, 나정조, “딥러닝 기반 수어 교육 온라인 플랫폼 구현”, *디지털콘텐츠학회논문지*, 23(11), 2022.11, 2,147-2,157.

반려견 감정 인식을 위한 CNN-LSTM 기반 이미지 캡션화 기술 개발

오승기, 백으뜸
 호남대학교 AI빅데이터학과
 e-mail : j030924@naver.com, geodol100@gmail.com

Developing a CNN-LSTM-Based Method for Captioning Emotions in Dog Images

Seung-Gi Oh and Eu-Tteum Baek
 Dept. of AI & Bigdata, Honam University

요약

본 논문은 이미지 캡션을 활용한 반려견 감정 인식 시스템을 제안한다. 제안된 시스템은 반려견 사진을 분석하여 감정을 식별하는 방식으로, 반려견의 비언어적 의사소통을 보다 효과적으로 이해하고 해석하는 데 중점을 둔다. 하지만, 반려견의 감정 상태를 정확하게 파악하는 것은 도전적인 과제이다. 이를 해결하기 위해, 본 연구는 행복, 편안함, 슬픔, 분노의 네 가지 주요 감정을 감지할 수 있는 이미지 캡션 모델을 설계하였다. 각 감정에 대해 1,000개의 샘플 데이터를 수집하여, 이를 설명적 텍스트로 변환하는 과정을 통해 감정 인식 데이터셋을 구축하였다. 수집된 반려견 감정 데이터는 먼저 사전 처리된 CNN 모델을 통해 이미지에서 주요 특징을 추출한다. 이후, 이 특징 정보는 사용자 지정 JSON 캡션 파일에 임베딩되어 LSTM 네트워크로 처리된다. 마지막으로, CNN과 LSTM의 결과를 병합하여 최종 감정 예측을 도출함으로써, 반려견의 감정 상태를 보다 정확하고 신뢰성 있게 분석할 수 있도록 한다.

1. 서론

이미지 캡셔닝은 시각적 이해 시스템과 의미 있고 구문적으로 올바른 문장을 생성할 수 있는 언어 모델을 사용하여 자연어로 이미지의 시각적 내용을 설명하는 작업이다[1]. 최근 신경과학의 발전으로 인간의 두뇌가 시각과 언어 생성을 어떻게 연결하는지에 대한 이해가 깊어지며 인공 지능 연구자들은 이러한 발전에 영감을 받아 이미지를 처리하고 해당 텍스트 설명을 생성할 수 있는 아키텍처를 개발해 왔다. 이러한 연구의 목표는 언어의 유창함을 유지하면서 입력 이미지를 처리하고, 그 내용을 표현하고, 시각적 요소와 텍스트 요소 간의 연결을 생성하여 단어 시퀀스로 변환하는 가장 효과적인 파이프라인을 찾는 것이다[2]. 본 논문에서는 이미지에 포착된 반려견의 감정 상태를 보다 정확하고 신뢰성 있게 분석하는 것을 목표로 하는 CNN-LSTM 이미지 캡셔닝 모델을 제안한다.

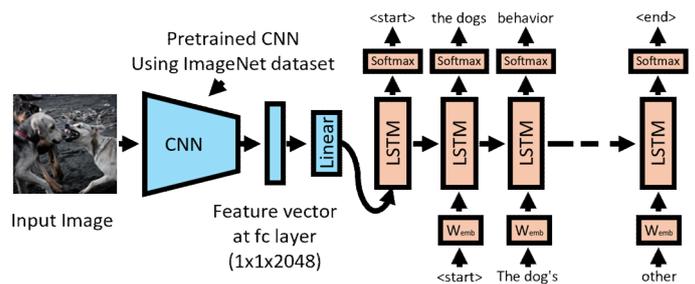
2. 관련 연구

2.1 CNN-LSTM

CNN-LSTM 이미지 캡셔닝 모델은 CNN의 강력한 이미지 특징 추출 기능과 LSTM 네트워크의 시퀀스 모델링 강점을 통합한다. 이 아키텍처에서 CNN 구성 요소는 일련의 convloutin 및 pooling 레이어를 통해 입력 이미지를 처리하여 공간 계층을 캡처하고 이미지의 풍부한 기능 표현을 생성한다. InceptionNet과 같은 고급 CNN 아키텍처는 세부적인 특징을 추출하는 데 있어 높은 정확성과 효율성으로 인해 종종 사용된다. 이미지 특징이 추출되면 LSTM 구성 요소에 대한 입력 역할을 하는 고정 길이 벡터로 평면화된다.

LSTM 네트워크는 평면화된 이미지 특징을 취하고 일련의 단

어를 생성하여 이미지에 대한 설명 캡션을 효과적으로 형성한다. 기존 순환 신경망(RNN)과 달리 LSTM에는 입력, 출력 및 망각 게이트와 같은 특수 게이트가 있어 장기적인 종속성을 처리하고 기울기 소실 문제와 같은 문제를 완화할 수 있다. 이로 인해 LSTM은 일관되고 문맥상 정확한 문장을 생성하는 데 특히 적합하다. 결합한 CNN-LSTM 아키텍처는 end-to-end 훈련을 사용하여 이미지 특징 추출과 시퀀스 생성 프로세스를 동시에 최적화한다[3][4]. CNN-LSTM 모델 구조는 그림 1과 같다.



(그림 1) CNN-LSTM 모델 구조

3. 실험 결과

3.1 데이터 전처리

Kaggle의 Dog Emotion 데이터 세트를 사용했으며, 데이터 세트에는 4가지 감정이 사용된 4,000개의 데이터가 들어 있다. 각 감정은 1,000개의 데이터를 가지고 있으며, 감정은 행복, 화남, 슬픔, 편안함이다. 본 논문에서는 표 1에서 볼 수 있듯이 Dog Emotion 데이터 세트에서 감정에 관한 여러 단어를 넣은 사용자 지정 JSON 파일을 만들어서 사용했다.

각 이미지와 해당 이미지의 캡션을 전처리한다. 전처리 과정은 다음과 같다. 각 이미지의 주석 파일을 가져온 후, 이미지와 캡션

을 로드하고, 로드된 캡션은 모든 구두점을 제거하고, 하이픈(-)을 공백으로 바꾸고, 캡션을 단어 단위로 나누고, 모든 단어를 소문자로 변환한다. 단어 목록은 다시 문장으로 결합하고, 각 캡션에 시작 태그(<start>)와 종료 태그(<end>)가 추가된다.

전처리 된 캡션은 키로 이미지 파일 이름을 사용하여 사전에 저장되었으며, 최종 데이터 세트는 이미지 파일 이름과 이미지에 대한 전처리 된 캡션 목록으로 구성된다. 그런 다음 캡션 텍스트를 전처리하여 이미지 캡션 생성 모델을 학습하고 토큰라이저를 사용하여 단어 인덱스를 생성한다. 모든 캡션을 단일 목록으로 평면화하고 각 이미지의 캡션 목록을 하나의 목록으로 결합하여 전체 데이터 세트의 모든 캡션을 포함한다.

Id	이미지	Caption
1		The dog is furious.
2		The dog is enraged.
1001		The dog is full of joy.
1002		The dog is wagging its tail.
2001		The dog is calmly lounging around.
2002		The dog's relaxed demeanor is soothing.
3001		The dog looks melancholic today.
3002		The dog appears downcast and somber.

(표 1) 이미지 캡셔닝 데이터

3.2 분석 결과

이미지를 299x299픽셀로 크기 조정하고 정규화하여 특징 벡터를 추출한다. 이를 통해 고차원 이미지를 효과적으로 학습할 수 있다. 캡션을 단어 시퀀스로 변환한 후 여러 개의 임베딩 쌍으로 분할하여 입력 시퀀스를 최대 길이로 채운 다음 원핫 인코딩 방식을 사용하여 출력 시퀀스를 변환한다. 이 CNN-LSTM 조합 모델의 경우 Description Length를 14로 설정하고 Vocabulary Size를 779로 설정했다. 2,048개의 필터에 50% Dropout을 적용하여 CNN 레이어의 설정을 256개의 필터로 줄인 다음 활성화 함수 ReLu를 적용하여 크기를 줄였다. 시퀀스 분류 문제에서 LSTM을 단방향으로 설정하고 임베딩 후 50% Dropout을 수행하여 256개의 뉴런을 생성했다. 그리고 디코더의 경우 CNN과 LSTM을 결합하여 최종 출력을 생성한다. CNN 모델은 InceptionV3 모델을 사용했다. 제안하는 기법의 정확성을 확인하기 위해 4가지 유형의

감정 데이터와 손실 값을 사용하여 이미지 캡셔닝이 성공했는지 확인했다. 제안하는 기법의 분석 결과는 <그림 2>에 나와 있다.



(그림 2) 이미지 캡셔닝 분석 결과

4. 결 론

본 논문에서는 딥러닝 알고리즘 CNN과 LSTM을 결합하여 Dog Emotion 데이터를 이용한 이미지 캡셔닝을 제안하였다. 본 연구의 문제점은 사진만 보고는 개의 감정을 인식하기 어렵다는 것이다. 또한 loss값이 어느 순간 하락이 진행되지 않는 것으로 보아 이는 백터의 차원이 지나치게 커질 수 있는 문제인 "차원의 저주"로 알려진 문제일 수 있음을 시사한다. "차원의 저주"는 데이터 집합을 늘리거나 학습률을 조정하면 해결할 수 있지만, 데이터 집합을 늘리면 더미 데이터나 사진 저작권 문제가 발생하고, 학습률을 조정하여 해결하기에는 캡셔닝이 진행이 되지 않는다. 본 연구 결과에 관한 향후 연구 방향은 추가 데이터 집합을 사용하거나 CNN 모델을 VIT 모델 또는 상위 모델로 변경하거나 CNN에 Video Transformer를 삽입하는 것이다.

참고 문헌

[1] A. Ardila, B. Bernal, and M. Rosselli, "Language and visual perception associations: meta-analytic connectivity modeling of Brodmann Area 37," Behavioural Neurology, 2015.

[2] Stefanini, M., Cornia, M., Baraldi, L., Cascianelli, S., Fiameni, G., & Cucchiara, R. (2022). From show to tell: A survey on deep learning-based image captioning. IEEE transactions on pattern analysis and machine intelligence, 45(1), 539-559.

[3] Park, H., & Kim, K., (2019). Sentiment Analysis of Movie Review Using Integrated CNN-LSTM Model, Korean Society of Intelligent Information Systems, 25(4), 141-154.

[4] Kim, Y., Lee, S., & Kwon, Y. (2020). Proposal of a step-by-step optimized campus power forecast model using CNN-LSTM deep learning. Journal of the Korea Academia-Industrial cooperation Society, 21(10), 8-15.

감사의 글

본 과제(결과물)는 과학기술정보통신부의 출연금으로 수행한 AI융합대학 지원사업 연구결과입니다.

맞춤형 디지털 리터러시를 위한 LLM과 지식 그래프를 활용한 프레임워크 설계

임가겸, 박정현, 한정규, 천세진
동아대학교 컴퓨터AI공학부

{1908313, pwjdgus1017}@donga.ac.kr, {jkhan, sjchun}@dau.ac.kr

Designing a Framework Using LLM and Knowledge Graph for Customized Digital Literacy

Kakyeom Lim, Jeonghyeon Park, Jungkyu Han, Sejin Chun
Department of Computer Engineering and AI, Dong-A University

요약

본 연구는 Large Language Model(LLM)과 Knowledge Graph를 사용하여 디지털 리터러시 증진을 위한 도구를 제안한다. 현재의 디지털 포용 정책을 비판하고, 소외계층을 동적으로 분류하여 맞춤형 지원을 제공하는 프레임워크를 제안한다. 제안된 프레임워크는 LLM을 통해 공공서비스 접근을 단순화하고, 사용자의 경험을 개선한다. 또한, Knowledge Graph 기반 접근 방식을 통해 사용자의 이해 수준에 맞춘 교육 콘텐츠를 제공하여 깊이 있는 학습과 재학습의 기회를 제공한다.

1. 서론

한국의 ICT 인프라는 전 세계적 관점에서 매우 우수하다고 평가받는다. 하지만 고령층으로 대표되는 디지털 소외계층의 디지털 격차는 경제협력개발기구 회원국 중 가장 미흡한 수준이다.

[1]에 따르면 디지털 격차는 오랫동안 사회·경제적 해결 과제로써 주지되어왔다. 하지만 오늘날에도 여전히 소외계층은 사회적 참여 및 평생학습 기회의 사각지대에서 벗어나지 못하고 있다.

이와 같은 격차를 해소하기 위한 디지털 리터러시는 기술적 사용 능력뿐만 아니라, 정보에 대한 접근성과 이를 비판적으로 이해하고 창의적으로 활용하는 능력을 포함하는 개념으로, 디지털 격차를 해소하는 핵심 요소다.

본 연구는 소외계층의 디지털 격차 해소를 위해 Large Language Model(LLM)과 Knowledge Graph 기반 디지털 리터러시 증진을 위한 도구를 제안한다.

2. 관련 연구

디지털 전환 과정에서 고령층과 소외계층은 디지털 리터러시의 부족과 기술에 대한 불안으로 인해 사회·경제적 불평등이 심화할 수 있다고 지적했다[2]. 특히, 공공기관의 디지털 포용 정책은 계층 간 차이를 충분히 반영하지 못하는 한계가 있다[3]. 한편, 디지털 학습 도구가 고령층의 불안감을 줄이고 자신감을 높여, 기술 활용 능력을 향상시킬 수 있다는 것을 밝혔다[4]. 이러한 연구는 정적인 인구통계학적 정보만 활용해 디지털 소외계층을 분류한다. 이는 코호트 효과를 고려하지 않아 디지털 리터러시 증진의 대상을 잘못 선정할 우려가 있다. 따라서, 디지털 소외계층을 동적으로 분류할 수 있는 새로운 방법이 필요하다.

3. 맞춤형 디지털 리터러시 증진 도구

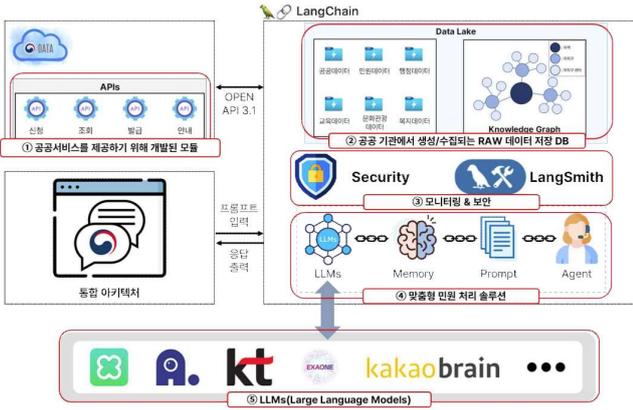
이번 장에서는, 소외계층의 디지털 리터러시 증진을 위해 LLM과 Knowledge Graph가 결합된 프레임워크를 설명한다.

3.1 LLM 기반 공공서비스 절차 간소화 프레임워크

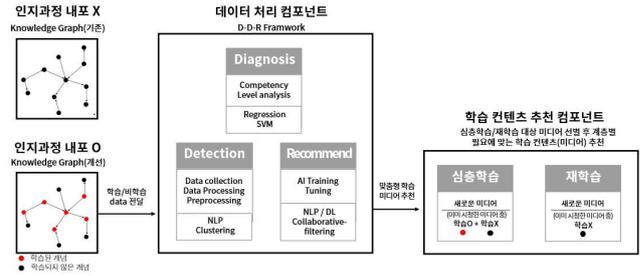
그림 1은 공공서비스 접근성 개선을 위해 LLM 기반의 통합된 API 사용 환경을 제공하는 프레임워크다. 우리는 그림에 기입한 번호 순서대로 각 모듈에 대해 설명한다.

본 프레임워크는 ①공공서비스 수혜를 위해 각기 접속 및 요청해야 하는 API들을 일괄적으로 조율한다. ②API 처리에 필요한 데이터의 조달/갱신을 위해 민간-공공 DB를 활용하고, Knowledge Graph를 통해 사용자 중심의 DB 조달을 프로비저닝한다. ③이렇게 관리되고 있는 데이터는 LLMops 도구인 LangSmith를 통해 모니터링 및 검토하고 보안을 유지한다. ④이후 사용자의 프롬프트 요청을 관련 부서 행정 처리에 파인 튜닝된 Agent가 처리하여, 인구통계학적 정보를 토대로 최적의 공공서비스를 중개한다. 이때 통합 아키텍처는 사용자 친화적 인터페이스를 제공한다. 이는, 기존에 직접 인터넷을 통해 검색 및 조회해야 했던 작업을 통합 처리하도록 한다. 정부 24로부터 공공서비스 수혜에 필요한 신청-조회-발급-안내에 대한 API를 통합적으로 가져온다.

제안하는 프레임워크는 ⑤LLM 기반의 통합된 User Interface를 통해 기존의 공공서비스의 복잡한 절차를 한번의 처리 과정으로 축약한다. 특히, memory 모듈은 사전에 전달된 사용자 정보를 활용한다. 이를 통해 공공서비스 이용 시 사용자 정보가 필요한 행정 처리시간이 축소된다.



(그림 1) LLM 기반 절차감소형 프레임워크



(그림 2) Knowledge Graph 기반 인지 단계별 프레임워크

3.2 Knowledge Graph 기반 학습 프레임워크

그림 2는 전통적으로 분류되어온 고령층, 저소득층, 이민자 등의 디지털 소외계층 뿐 아니라 다양한 디지털 숙련도를 지닌 사용자를 고려할 수 있는 프레임워크를 제시한다.

우선 기존의 Knowledge Graph를 개선하여 학습자료를 경험한 후 아직 이해하지 못한 개념의 노드와 이해한 개념의 노드를 구분한다. 이를 통해 단지 한 번 경험했을 뿐인데 이해를 끝마친 것으로 간주되는 기존 Knowledge Graph의 문제를 보완한다.

인지과정이 내포된 노드가 데이터 처리 컴포넌트에 전달되면 Detection-Diagnosis-Recommend(DDR) Process를 통해 사용자의 학습 데이터가 처리되어 맞춤형 영상을 추천한다. DDR 프레임워크의 Detection 모듈은 NLP Clustering을 통해 Knowledge의 유사도를 기반으로 군집화를 진행한다. 이후 Diagnosis 모듈에서 Support Vector Machine(SVM) 회귀분석을 활용하여 인지 단계를 구분/분석한다. 분석된 인지 단계를 기반으로 Recommend 모듈에서 Collaborate Filtering 추천 시스템을 활용하여 사용자에 맞춤형 학습자료를 추천한다.

이때 학습 추천 컴포넌트에서는 심층학습과 재학습 자료를 하이브리드로 추천한다. 심층학습을 위한 콘텐츠의 경우, 학습한 자료들에 대해 아직 이해하지 못한 개념을 기반으로 새로운 학습자료를 추천한다. 또한 재학습을 위한 콘텐츠의 경우, 학습한 자료들에 대해 이해를 완료한 개념을 기반으로 새로운 학습자료를 추천한다. 이를 통해 사용자의 인지 레벨에 맞춤형 학습 자료를 추천하여, 단지 오늘날 기술 수준을 위한 학습 보조 도구 뿐 아니라 빠르게 발전하는 신기술 생태계에서 지속적으로 활용할 수 있는 프레임워크를 제시한다.

4. 결론 및 향후 연구

본 연구는 LLM과 Knowledge Graph 기반 프레임워크를 통해 소외계층의 디지털 접근성과 디지털 리터러시 격차 문제를 해결하고자 했다. 이때 같은 연령이라도 세대별 문화 특성(산업화세대, 베이비붐세대, X세대 등)에 따라 다양한 디지털 숙련도를 고려할 수 있는 프레임워크를 제시하여, 향후 도구 개발에 있어 코호트 효과를 간과하는

오류를 방지한다는 데에 의의가 있다.

또한 빠르게 발전하는 기술의 추세를 고려하여, 기존에 취약계층으로 분류되던 사람들 이외의 새로운 디지털 소외계층을 포용할 수 있는 프레임워크를 제시하였다. 이에 단순히 ‘소외계층’으로 포괄되는 집단이기보다, 더욱 섬세화된 계층·사회 집단을 체계적으로 고려할 수 있는 프레임워크로 확대 개발해가고자 한다.

감사의 글

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가의 지원을 받아 수행된 연구임 (No. 2023-0-00076, SW중심대학(동아대학교))

참고 문헌

- [1] Peng, D., & Yu, Z. (2022). A literature review of digital literacy over two decades. *Education Research International*, 2022(1), 2533413.
- [2] Vassilakopoulou, P., & Hustad, E. (2023). Bridging digital divides: A literature review and research agenda for information systems research. *Information Systems Frontiers*, 25(3), 955-969.
- [3] 배영임 (2022). A Study on Inclusive Growth Policy from the Perspective of Technology Innovation: Focusing on Digital Inclusion. Gyeonggi Research Institute.
- [4] Martínez-Alcalá, C. I., Rosales-Lagarde, A., Pérez-Pérez, Y. M., Lopez-Noguerola, J. S., Bautista-Díaz, M. L., & Agis-Juarez, R. A. (2021, July). The effects of Covid-19 on the digital literacy of the elderly: norms for digital inclusion. In *Frontiers in Education* (Vol. 6, p. 716025). Frontiers.

OTT 기반의 모션 검출 및 객체인식 모델 연구

김유림¹, 이승재¹, 최희녕¹
전북대학교 IT정보공학과

e-mail : yoo9379@gmail.com, leesengjie@naver.com, soonso00@naver.com

A Study on OTT-based Motion Detection and Object Recognition Models

Yoo-Rim Kim, Seung-Jae Lee, Hwi-Nyeong Choi
Dept. of IT Information Engineering, Jeonbuk National University

요약

최근 OTT 서비스의 이용 증가로 인해 많은 사람들이 일방향 미디어에 장시간 노출되고 있다[1]. 이는 영유아의 발달 과정에 부정적인 영향을 미치게 된다. 따라서 본 논문은 영유아의 건강한 미디어 매체의 접근을 위해 영유아가 화면 속 동작을 따라 움직이며 학습에 능동적으로 참여할 수 있는 상호작용 학습 콘텐츠를 개발하는 것을 목적으로 한다. OTT 셋톱박스에 내장된 카메라를 활용해 사용자 모션 인식을 진행하여, 영유아가 콘텐츠에서 요구하는 행동을 올바르게 수행하는지를 판단하는 기능을 구현하였다. 사용자 모션 인식은 MediaPipe Holistic을 활용하여 실시간으로 움직임을 감지하고, 학습에 사용되는 교구재 인식을 위한 객체 인식은 Custom CNN 모델을 통해 그의 사용 여부를 판단할 수 있게 제안했으며, 본 논문에서 객체 인식 모델은 85.71%의 정확도를 구현하였다.

1. 서론

코로나(COVID-19) 사태 이후 OTT(Over-The-Top) 서비스에 대한 수요가 크게 급증해 콘텐츠 소비 패턴이 재편되었다. 연구에 따르면 봉쇄 기간 동안 더 많은 사람들이 집에 머무르며 엔터테인먼트를 위해 스트리밍으로 전환하면서 OTT 플랫폼 사용과 기존 TV 시청률이 모두 증가했다[1].

하지만 뇌 발달이 중요한 시기인 영유아기에 장시간 동안 일방향 미디어 매체에 노출되는 것은 그들의 발달 과정에 부정적인 영향을 미칠 수 있다[2]. 2015년 옥스퍼드 대학의 ‘텔레비전 시청이 뇌구조에 미치는 영향: 횡단면 및 종단면 분석’ 논문에서 따르면 TV 시청이 전두엽을 포함한 특정 뇌 영역의 회백질 및 백질 구조에 변화를 일으켜 언어 지능(IQ) 저하에 영향을 미칠 수 있다고 보고되었다[3]. 전두엽은 결정을 내리고, 계획을 세우고, 사회적 상호작용을 처리하고, 감정을 조절하는 등의 중요한 기능을 담당한다.

그러나 미디어 매체의 인기와 사용량이 증가함에 따라 현대인의 일상에서 TV는 필수불가결한 도구로 기능한다. TV와 같은 일방향 미디어의 한계를 극복하고 전두엽 발달을 촉진하기 위해, 사용자가 능동적으로 반응하고 참여할 수 있는 상호작용 가능한 뉴미디어의 필요성이 더욱 강조되고 있다.

따라서 본 논문은 게임 콘솔인 Nintendo Wii와 같이 사용자와 상호작용할 수 있는 학습 콘텐츠를 개발하기 위한 목적으로 연구를 진행한다. 이러한 상호작용 학습 콘텐츠는 사용자가 화면 속 동작을 따라 하며 자신의 신체를 움직이고, 이를 통해 능동적으로 학습에 참여하도록 설계된

다. 모션 인식 기술을 활용해 사용자의 동작을 실시간으로 추적하고 분석함으로써, 상호작용을 유도하는 방식이다.

MediaPipe Holistic을 활용해 사용자의 움직임을 효과적으로 인식하고, 사용자의 몸에 Bounding Box를 처리해 범위 내에서 특정 동작의 수행 여부를 판단한다. 더 나아가 Custom CNN을 통해 객체 분류 기술을 구현해 학습에 사용되는 교구재를 인식할 수 있도록 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 사용자 모션 인식과 특정 행동 수행 여부 판단, 그리고 특정 교구재 객체 인식에 대해 나누어 설명한다. 마지막으로 3장에서 결론과 향후 연구방향으로 맺는다.

2. 관련 연구

2.1 사용자 모션 인식 및 특정 행동 수행 여부 판단

MediaPipe 솔루션은 구글에서 제공하는 인공지능(AI) 및 머신러닝(ML) 기법을 빠르게 적용할 수 있는 라이브러리 서비스이다. 그중 MediaPipe-Holistic은 기존 Mediapipe의 Face-Mesh, Pose, Multi-hand를 합쳐놓은 모듈로 구성되어 있다. Face-Mesh의 경우 468개의 3D 얼굴 랜드마크를 실시간으로 추정하는 얼굴 형상 솔루션이다. Pose는 고충실도 신체 자세 추적을 위한 ML 솔루션으로, BlazePose 연구를 사용하여 RGB 비디오 프레임에서 전신의 33개의 3D 랜드마크를 추론한다. 마지막으로 Hands는 충실도가 높은 손 및 손가락 추적 솔루션으로, ML을 사용하여 단일 프레임에서 손의 3D 랜드마크 21개를 추론한다.

본 논문은 MediaPipe Holistic을 사용하여 사용자 모션 인식을 구현하였으며, 랜드마크 시각화를 위해 MediaPipe의 drawing_utils를 사용하고, 그림 1의 좌와 같이 각각의 랜드마크를 선으로 이어 화면에 나타내었다.

* 제 1저자: 학생 회원 김유림(전북대학교 IT정보공학과), 공동 저자: 이승재(전북대학교 IT정보공학과), 최희녕(전북대학교 IT정보공학과)

** "본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 지원을 받아 수행"

특정 행동의 정답 데이터를 이미지 형태로 입력한 후 그림 1의 좌와 같이 바운딩 박스를 둘러싼다. 정답 데이터의 좌표값을 바운딩 박스 기준으로 정규화 시켜 이후 비디오의 랜드마크 값과 비교할 수 있도록 한다. 정답 이미지와 실시간 모션이 인식되는 프레임 속 사람에게 각각 바운딩 박스를 생성하고, 그 랜드마크를 바운딩 박스를 기준으로 정규화시킨다. 그리고 두 개의 3D 랜드마크 좌표값 집합을 비교해 일정 임계값 내에서 일치하는 점들의 개수를 집계했다. 이때 임계값이란, 두 랜드마크의 좌표가 일정 거리를 벗어나지 않은 범위에서 일치하는 것으로 간주하는 기준값으로, 이 과정에서 표 1과 같이 임계값은 0.05로 설정하였다. 이는 유아의 특성상 오랫동안 동일한 행동을 유지하기 어렵다는 점을 고려해, 정확한 좌표로부터 일정 수준의 오차 값을 허용한 상태로 앞서 추출했던 랜드마크 값을 활용해 특정 행동 수행 여부 판단을 진행하였다.

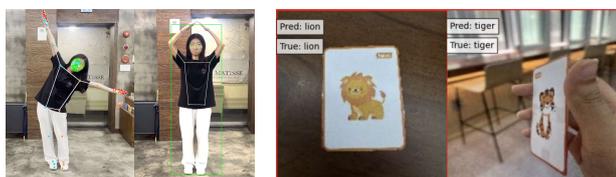
(표 1). 임계값 수식 관한 함수

```
def compare_landmarks(landmarks1, landmarks2,
threshold=0.05):
    matching_points = 0
    for lm1, lm2 in zip(landmarks1, landmarks2):
        if abs(lm1['x'] - lm2['x']) <= threshold and
abs(lm1['y'] - lm2['y']) <= threshold and
abs(lm1['z'] - lm2['z']) <= threshold:
            matching_points += 1
    return matching_points
```

2.2 특정 교구재 객체 인식

특정 교구재를 인식하기 위한 맞춤형 CNN(Convolutional Neural Network) 모델을 설계하고, 이를 학습시키기 위해 제작한 교구재 데이터셋을 사용했다. 주 교구재로는 그림1에 우와 같이 사자와 호랑이 그림이 그려진 카드를 활용했으며, 두 객체를 정확히 분류하는 모델을 구현했다. CNN 모델은 두 개의 Convolution layer와 Pooling layer, 그리고 두 개의 Fully connected layer를 기반으로 설계했다.

데이터셋은 그림 1의 우와 같이 사자와 호랑이 카드를 각각 클래스 0과 1로 라벨링하여, 약 70%의 데이터를 Train data set으로, 나머지 30%를 Validation data와 Test data로 각각 나누었다. 데이터셋이 비교적 작다고 판단하여, Resizing한 이미지 파일을 Blur, 밝기 조정, 이미지 축소, 좌우 반전을 적용시켜 기존 데이터를 7배로 증강시켰다. 이외에도 훈련 중 Overfitting을 방지하기 위해 Dropout[4]과 Early Stopping 기법을 적용하였다. 이러한 기법들을 적용한 결과, CNN 모델은 Test set에서 85.71%의 정확도를 가졌으며, 이는 교구재 객체 인식 모델로서 실용적인 성능을 보인다고 판단된다.



(그림 1). 모션 인식 및 교구재 객체 인식

3. 결론

본 논문에서는 미디어가 일상화된 현대사회에서 영유아의 건강한 미디어 매체의 접근을 유도하고 능동적인 참여를 요구하는 방법을 연구하였다. 영유아 대상으로 체육 및 요가 등의 분야에서 각 개인의 특성과 요구에 맞춘 차별화된 서비스를 제공하고 인지, 운동, 사회성 발달 등의 발달 과업을 지원하며 이를 통해 일방향 미디어 문제를 효과적으로 해결할 수 있기를 기대한다. 결론적으로 본 논문에서는 상호작용할 수 있는 미디어 개발을 위해 MediaPipe Holistic을 활용하여 영유아 모션 인식을 구현하고, 실시간으로 영유아의 동작을 추적하여 특정 행동 수행 여부를 판단할 수 있는 시스템을 개발하였다. 또한, Custom CNN 모델을 통해 교구재 인식을 수행하였으며, 도출된 객체 인식 모델의 정확도는 85.71% 정도의 결과를 보였으며, 향후 연구에서 학습 데이터셋을 더 확보하여 정확도를 높이는 연구를 진행할 예정이다.

참고 문헌

[1] Rajani, M., & Rajani, S., "A study on adoption of OTT platforms and its growth in the post-COVID era", Webology, Vol.18 No.6, 2021. pp.417. DOI: <http://www.webology.org>

[2] Heffler, K. F., Sienko, D. M., Subedi, K., McCann, K. A., & Bennett, D. S., "Association of Early-Life Social and Digital Media Experiences With Development of Autism Spectrum Disorder - Like Symptoms," JAMA Pediatrics, Vol.174 No.7, 2020, pp.690-696. DOI: <https://doi.org/10.1001/jamapediatrics.2020.0230>

[3] Takeuchi, H., Taki, Y., Hashizume, H., Asano, K., Asano, M., Sassa, Y., ... & Kawashima, R., "The impact of television viewing on brain structures: Cross-sectional and longitudinal analyses", Cerebral Cortex, Vol.25 No.5, 2015. pp.1188. DOI: <https://doi.org/10.1093/cercor/bht315>

[4] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R., "Dropout: A simple way to prevent neural networks from overfitting", Journal of Machine Learning Research, Vol.15 No.56, 2014. pp.193-1931.

YOLOv8 기반 실시간 태양광 패널 오염물질 인식 및 자동 청소 로봇

송상진, 민석주, 정희자**, 최광미, 김남호*

*호남대학교 컴퓨터공학과, **(주)휴넷가이아

ssj30099@naver.com, cherry99goo@gmail.com, jeong008@nate.com,
{cgmi66,*nhkim}@honam.ac.kr

YOLOv8-based real-time solar panel contamination and automatic cleaning robot

Sangjin Song, Seokju Min, Heeja Jeong, Gwangmi Choi, Namho Kim*

*Dept. of Computer Engineering, Honam Univ. **Hunetgaia Co., Ltd.

요약

본 연구에서는 태양광 패널의 성능 저하를 초래하는 다양한 오염물질을 실시간으로 탐지하고, 오염된 부분만을 선택적으로 청소하는 로봇 시스템을 개발하였다. 약 2500장의 이미지를 활용해 YOLOv8 모델을 학습시켰으며, 이 모델을 통해 태양광 패널 위의 오염 상태를 정확하게 인식할 수 있도록 하였다. 로봇은 메카넘 휠을 사용해 복잡한 패널 배열에서도 효율적으로 이동하며, 라즈베리파이 4와 아두이노를 결합해 청소 작업을 수행한다. 본 연구는 태양광 패널의 유지관리 효율성을 높이고, 자원의 낭비를 줄이며, 경제적인 운영을 가능하게 하는데 중요한 기여를 할 것으로 기대된다.

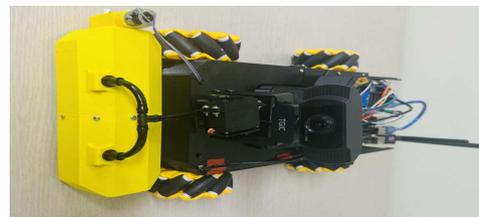
1. 서론

지속 가능한 에너지 자원의 중요성이 커짐에 따라 태양광 발전이 전 세계적으로 중요한 에너지원으로 부상하고 있다[1]. 태양광 패널의 효율은 흙, 눈, 먼지, 새똥 등의 오염물질에 의해 크게 저하될 수 있으며, 이는 패널 표면의 빛 흡수 능력을 방해하여 전력 생산 효율을 감소시킨다[2]. 따라서 태양광 패널의 정기적이고 효과적인 청소는 필수적이다. 전통적인 수동 청소 방식은 노동 집약적이고 비용이 많이 들며, 특히 넓은 면적의 패널이나 접근이 어려운 위치에 설치된 패널에서는 안전 문제를 초래할 수 있다[3]. 또한, 일괄 청소는 비효율적일 뿐만 아니라 패널 손상의 위험을 증가시킬 수 있다. 최근 연구들은 오염된 부분을 정확히 인식하여 청소하는 자동화 시스템 개발에 주목하고 있다[4][5]. 본 연구에서는 딥러닝 기반 객체 인식 모델 YOLOv8을 사용하여 태양광 패널 표면의 특정 오염물질을 실시간으로 탐지하고, 오염된 부분만을 중점적으로 청소하는 로봇 시스템을 제안한다. 이 로봇은 카메라를 통해 패널 표면을 모니터링하며 실시간으로 오염된 지점을 감지하고 선택적으로 청소를 수행하여, 불필요한 청소를 줄이고 청소에 소요되는 물과 에너지를 절약함으로써 태양광 발전의 효율성과 경제성을 극대화할 수 있는 효과적인 솔루션을 제공한다.

2. 태양광 패널 로봇 하드웨어 구성

본 연구에서 개발된 태양광 패널 청소 로봇은 실시간 오염물질 감지와 AI로 인식한 오염물질 부분의 선택적 청소를 위해 다양한 하드웨어와 소프트웨어 요소를 통합하

여 설계되었다. 이 로봇은 오염된 패널 부분을 정확히 탐지하고 오염된 부분을 중점적으로 청소함으로써 효율성과 비용 절감을 극대화하는 것을 목표로 한다. 로봇의 이동 방향과 진환에 있어서 메카넘 휠을 사용하여 45도 각도로 장착된 여러 개의 작은 롤러를 통해 360도 방향으로 자유롭게 움직일 수 있어 복잡한 패널 배열에서도 효율적으로 이동한다. 이를 통해 로봇은 오염된 지점을 신속하게 탐색하고, 정확한 위치에 접근하여 청소 작업을 수행한다. 3D 프린팅 기술로 제작된 로봇 본체는 경량화와 내구성을 확보하고, 구동 및 제어는 Arduino와 Raspberry Pi 4를 활용한다. Arduino는 모터, 브러쉬, 물 분사 시스템을 제어하며, Raspberry Pi 4는 YOLOv8 모델을 구동하여 실시간 오염물질 인식을 수행한다. 카메라로부터 받은 영상을 처리하고, Arduino에 명령을 전달하여 청소 작업을 수행한다. 운영체제로는 경량 리눅스 배포판인 Debian을 사용하여 소프트웨어의 안정성과 성능을 최적화한다.



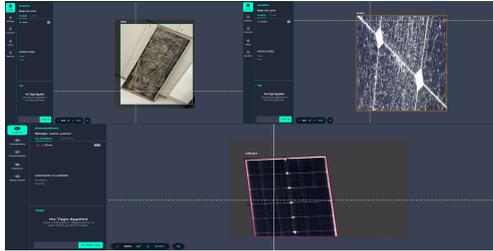
(그림 1) 태양광 패널 자동 청소로봇 외형

3. 시스템 구현

3.1 데이터 수집 및 라벨링

데이터세트는 태양광 패널위의 정상적인 패널인 Clean, 오염 Dirty, 결합 broken으로 3개의 클래스로 구분하였으

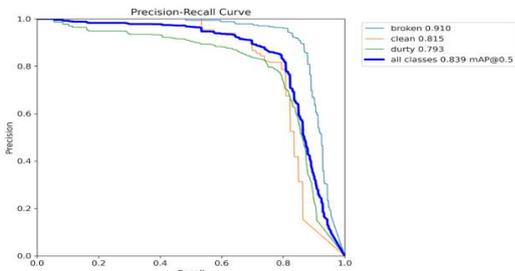
며 총 2273장으로 각 클래스마다 약 757장의 이미지 데이터셋으로 구성하였다. 모든 데이터셋은 Robotflow를 사용하여 라벨링을 진행하고 Train, Valid 그리고 Test로 나누어 구성하였으며 7:2:1 비율로 데이터를 분할하였다. 이미지 사이는 640*640으로 통일하여 라벨링을 진행한다.



(그림 2) 데이터 라벨링 과정

3.2 yolov8 학습

본 연구에서는 라즈베리 파이의 제한된 연산 능력과 자원 제약으로 인해 직접 학습을 진행하기보다는, 더 강력한 컴퓨팅 자원을 제공하는 구글 코랩에서 YOLOv8을 모델을 이용하며, 사이즈는 640*640, Batch 32, epoch 100으로 학습을 진행하였다. Adam 옵티마이저를 통해 모델의 가중치를 최적화했으며, 데이터셋은 7:2:1 비율로 학습, 검증, 테스트 세트로 분할한다. 그림 3은 모델 학습 결과인 정밀도와 재현율 사이의 관계를 나타내는 PR 곡선이다. broken, clean, dirty 클래스에 대해 각각 0.910, 0.815, 0.793의 평균 정밀도를 기록했으며, 전체 클래스에 대한 mAP50는 0.839로 나타났다. 이러한 결과는 모델이 태양광 패널의 다양한 상태를 효과적으로 감지할 수 있음을 보여준다.



(그림 3) YOLOv8 mAP50

실제 태양광 패널 인식 테스트 결과, 모델은 패널의 상태를 정확하게 인식하여 오염, 정상, 결함 상태로 분류하고, 각각의 상태에 대해 높은 신뢰도를 나타낸다. 오염 상태의 패널을 0.94의 높은 확률로 탐지하고, 정상 상태와 결함 상태의 패널도 각각 0.90, 0.96의 높은 확률로 정확히 분류한다.



(그림 4) YOLOv8 테스트 결과

4. 결론

본 연구에서는 태양광 패널의 오염물질을 실시간으로 탐지하고, 선택적으로 청소할 수 있는 자동화 로봇 시스템을 개발하여 태양광 발전의 효율성을 극대화하는 데 중점을 두었다. 약 2500장의 오염물질 이미지 데이터를 활용해 학습된 YOLOv8 모델은 다양한 오염 상태를 정확하게 감지할 수 있으며, 라즈베리파이 4와 아두이노의 결합을 통해 실시간으로 청소 명령을 수행한다. 특히 메카넘 휠을 사용한 로봇의 이동 시스템은 복잡한 패널 배열에서도 효율적인 이동을 가능하게 하여, 카메라로 오염물질을 인식한 부분을 중점적으로 정밀하게 청소할 수 있게 한다. 이러한 접근 방식은 태양광 패널의 유지보수에 필요한 노동력과 비용을 크게 절감할 수 있을 뿐만 아니라, 청소 과정에서의 물과 에너지 사용을 최소화하여 환경적 지속 가능성을 확보할 수 있다. 추후 연구에서는 데이터셋을 증가시켜 모델의 정확도를 향상시키고 YOLOv8 모델을 경량화시키기 위해 TensorFlow Lite로 변환하여 로봇에 탑재함으로써 처리 속도를 향상시킬 예정이다. 결론적으로, 본 연구는 태양광 발전 시스템의 성능을 유지하고, 장기적인 경제성을 강화하며, 지속 가능한 에너지 자원의 관리와 최적화에 기여할 수 있는 중요한 기술적 기반을 제공하였다. 이는 향후 태양광 발전 산업의 발전 방향에 중요한 시사점을 제공하며, 자동화된 청소 로봇의 상용화 가능성을 제시한다.

감사의 글

“본 과제(결과물)는 과학기술정보통신부의 출연금으로 수행한 AI융합대학 지원사업 연구결과입니다.”

참고 문헌

[1] Ali O M Maka, Jamal M Alabid, Solar energy technology and its roles in sustainable development, Clean Energy, Volume 6, Issue 3, June 2022, Pages 476 - 483.

[2] Michael L. Adekanbi, Ezekiel S. Alaba, Toluwalope J. John, Tomi D. Tundealao, Titilope I. Banji, “Soiling loss in solar systems: A review of its effect on solar energy efficiency and mitigation techniques,” Cleaner Energy Systems, Volume 7, 2024, 100094, ISSN 2772-7831.

[3] Bakhiyi B, Labrèche F, Zayed J. The photovoltaic industry on the path to a sustainable future—environmental and occupational health issues. Environment International. 2014 Dec;73:224 - 234.

[4] 이동구. “태양광 발전소 유지보수를 위한 태양광 패널의 오염인식 및 추적용 드론기반 딥러닝 모델 개발.” 국내 박사학위논문 성균관대학교 일반대학원, 2022. 서울.

[5] 한호상, 서장원. (2024). 드론 촬영 이미지와 딥러닝 기법을 이용한 태양광 패널의 먼지오염 탐지. 한국자원공학 회지, 61(1), 23-32.

자연어기반 맞춤형 도서 추천을 위한 생성형AI 사서

류동훈¹, 박소윤², 김성은³, 황두현⁴, 공지훈⁵,

김기두⁶, 김장환⁷, 서채연⁸, 김영철^{9*}

^{1,2,3,4}홍익대학교 소프트웨어융합학과,

⁵툰스퀘어,

⁶한국정보통신기술협회,

^{7,8,9*}홍익대학교 소프트웨어융합학과

{ryudong12¹,syp4560²,gnh-08088³}@naver.com, sosyjin@gmail.com⁴,

go400s@gmail.com⁵, kdkim@tta.or.kr⁶,

{lentoconstante⁷, chaeyun⁸, bob^{9*}}@hongik.ac.kr

Generative AI Librarian for Natural Language-based Customized Book Recommendation

Ryu Dong-hoon¹, Park So-yoon², Kim Seong-eun³, Hwang Du-hyeon⁴, Kong Ji Hoon⁵,

Kidu Kim⁶, Janghwan Kim⁷, Chaeyun Seo⁸, R. Young Chul Kim^{9*}

^{1,2,3,4}Dept. of Software and Communications Engineering, Hongik University,

⁵Toonsqre Corp.,

⁶Telecommunications Technology Association,

^{7,8,9*}Dept. of Software and Communications Engineering, Hongik University

요약

현재 도서 인구가 급격히 줄고, 수많은 도서가 출판되어 독자가 원하는 도서를 찾아 읽기도 어려운 실정이다. 이런 문제를 해결하기 위해, 생성형 AI 도서 추천이 바쁜 현대인에게 적절한 맞춤형 도서 추천 AI를 제안한다. 그를 통해 독자가 책을 찾고 읽는 과정에서 겪는 다양한 불편을 해결하기 위한 서비스를 목적으로 한다. 첫째, 간략한 줄거리만 기억하고 책 제목을 떠올리지 못하는 경우, 독자가 원하는 책을 쉽게 찾을 수 있도록 돕는 서비스를 한다. 둘째, 책을 사거나 대출하기까지의 복잡한 과정을 단순화하기 위해, 서점이나 도서관의 서적 현황을 제공하는 서비스를 한다. 이를 AI 도서 추천 서비스를 통해 현대인들의 도서 접근성 및 문해력을 높여 인간 삶의 질과 관조 할 수 있는 삶을 기대한다.

키워드: 도서, 생성형AI, 도서추천, AI서비스

1. 서론

독서는 우리 삶에 다양한 방식으로 긍정적인 영향을 미친다. 언어발달, 사고력 신장, 창의성 증진 등 다양한 인지 능력을 향상시켜 지적 능력을 키운다. 스트레스 해소와 심리적 안정에도 도움을 주어 정신적 건강을 증진 해주기도 한다. 또한 책을 읽으며 자신의 취향과 가치관을 발견할 수 있다. 책을 통해 다양한 관점과 경험을 접할 수 있어 사회적 유대감을 강화할 수 있다. 독서는 단순한 취미가 아닌 자아 발견 및 발전의 중요한 통로가 된다. 또한, 독서 부족이 지속된다면 학업 성취도 저하, 교육 격차 심화, 독서 문화 침체와 같은 개인의 성장과 사회 전반에 걸쳐 다양한 부정적 영향을 미칠 수 있다. 개인, 가정, 학교, 정부 등 사회 구성원 모두가 협력하여 독서 문화 활성화를 위해 노력할 필요가 있다. 도서 인구를 증가시키기 위해, 원하는 책을 쉽게 검색하여 독서할 수 있게 하고, 사용자가 기억하는 책의 정보, 어림פות이 기억나는 줄거리, 조연의 이름을 입력할 수 있는 웹사이트를 개발한다. 웹 사이트에서 생성형 AI를 통해 책을 검색하고, 가장 유사한 책들을 사용자에게 제공한다. 검색된 책들은 제목, 작가, 출판사 등의 정보와 함께 사용자에게 표시한다. 사용자는 해당 책들을 클릭하여 자세한 정보를 확인할 수 있다. 또한, 찾고자 하는 책의 가까운 서점이나 도서관에서 해당 책을 찾을 수 있는 정보를 제

공한다. 독자들이 잊어버린 책을 찾는 데 도움을 주어 독서 경험을 더욱 풍부하게 만든다. 사용자의 편의를 고려하여 책을 구매하거나 대여할 방법을 제공하여 독서를 즐기는 과정을 편리하게 한다. 사용자들 사이에서 책에 관한 대화와 공유를 유도하여 독서 커뮤니티를 활성화한다. 2장에서는 책을 구매하고, 읽을 수 있는 웹사이트를 언급하고, 3장은 생성형AI 추천 메커니즘에 대해 언급하고, 4장은 설계 및 구현, 5장은 결론을 언급한다.

2. 관련 연구

원하는 도서를 찾을 방법은 다양하다. 대표적으로 알라딘중고서점이 있다. 책 전문 중고거래 플랫폼 중 하나인 '알라딘중고서점'에서 원하는 책을 검색으로 찾을 수 있다 [1]. 다른 방법은 경기도 사이버 도서관과 같은 지역 내에서의 공공도서관에서 도서에 대한 자료 검색이 가능하다 [2].

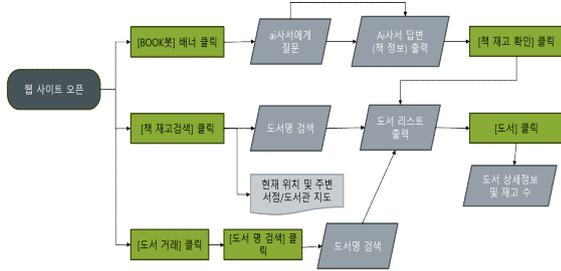
3. 생성형 AI 도서 추천 메커니즘

제안한 생성형AI 도서 추천 메커니즘은 책의 내용이나 등장인물을 토대로 책의 제목을 알려주고, 내가 원하는 카테고리의 책을 추천해 주는 BOOK BOT과 지도에서 책 제목을 입력하면 오프라인 매장의 책 재고 수를 지도로 한눈에 파악 할 수 있고, 오프라인 매장의 위치를 알 수 있는 BOOK MAP서비스 기능이다. 주요 기능은 총 2가지로, 생성형AI GPT-4-turbo를 활용한 도서 내용 및

상세 정보를 추천받을 수 있는 챗봇과 Kakao Search API를 활용한 도서 정보 API, 재고DB를 통해 도서 정보 검색 및 재고 확인을 할 수 있는 도서 정보 검색 기능이 있다 [3].

4. 설계 및 구현

생성형 AI기반 웹의 전체 설계 흐름도이다.



(그림 1) 전체 설계 흐름도

그림 1은 생성형 AI기반 웹의 전체 설계 흐름도이다. 웹사이트에서 BOOK봇 배너를 클릭하면 AI사서에게 질문하는 북봇이 나오고 AI사서는 책 정보를 출력한다. 책 재고 확인 버튼을 누르면 질문한 도서명의 리스트가 출력되고 해당 도서를 클릭하면 도서 상세 정보와 재고 수를 알 수 있다. 웹사이트에서 책 재고 검색을 누르면 현재 위치와 주변 서점과 도서관 위치의 지도가 나오고 도서명을 검색할 수 있다. 검색 후, 도서 리스트가 출력되고 원하는 도서를 클릭하면, 도서 상세 정보와 재고 수를 알 수 있다. 웹사이트에서 도서 거래를 클릭 후 나오는 배너의 도서명 검색을 클릭하면 책 재고 검색과 마찬가지로 도서명 검색 페이지 나와 도서 리스트 출력이 가능하고 원하는 도서를 클릭 시 도서 상세 정보와 재고를 알 수 있다.

메인 페이지에서 책 재고 검색을 클릭 시 재고 확인 페이지로 이동하고, 내가 원하는 책 검색 Book봇 이라는 배너를 클릭하면 Book Bot으로 이동한다. 그림 2는 BOOK BOT 페이지이다. BookBot 페이지에서는 “AI사서에게 질문해 보세요!” 라고 적혀 있는 입력 칸에 도서 내용이나 등장인물 및 추천받고 싶은 책을 입력한다. 도서 거래를 클릭하면 도서명 검색 배너가 나오고 클릭하면 도서명 검색 페이지로 넘어간다. 입력한 내용에 맞는 도서 정보가 제공되고 하단에는 도서 이미지와 상세 정보가 같이 제공한다.



(그림 2) BOOK BOT 페이지

이 페이지는 메인 페이지에서 책 재고 검색 버튼을 클릭하거나 BookBot의 결과 화면 페이지 책 재고 확인 버튼을 클릭하면 나오는 페이지로 오른쪽 지도에서는 나의 현재 위치를 알 수 있고 주변에 서점 및 도서관 위치를 확인 할 수 있다. 왼쪽은 책 제목 검색을 할 수 있는 칸이 있다. 책 제목 검색 칸에 제목을 검색하면 도서 리스트가 나오고 리스트에 나온 도서를 클릭 시 상세 정보

와 재고 수를 출력한다. 지도에서는 도서의 재고가 있는 지점을 마커로 표시한다. 도서를 클릭 시 나오는 배너에서는 상세한 책 정보와 서점의 재고 수량을 출력하여 도서 현황을 한눈에 볼 수 있다.

그림 3은 BOOT BOT 실행 페이지이다.



(그림 3) BOOK BOT 실행 페이지

그림 4는 서점/도서관 지도 페이지이다.



(그림 4) 서점/도서관 지도 페이지

5. 결론

현재 독서 인구가 급격히 줄고, 수많은 도서가 출판되어 독자가 원하는 도서를 찾아 읽기도 어려운 실정이다. 이런 문제를 해결하기 위해, 생성형 AI 도서 추천이 바쁜 현대인에게 적절한 맞춤형 도서 추천 AI를 제안한다. 그를 통해 독자가 책을 찾고 읽는 과정에서 겪는 다양한 불편을 해결하기 위한 서비스를 목적으로 한다. 즉, AI 도서 추천 서비스를 통해 현대인들의 독서 접근성 및 문해력을 높여 인간 삶의 질과 관조 할 수 있는 삶을 기대한다.

ACKNOWLEDGMENT

본 연구는 2024년도 문화체육 관광부의 재원으로 한국콘텐츠진흥원 (과제명: 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술개발, 과제번호: RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임. 또한, 2024년 1학기 홍익대학교 메타버스 융합SW 아카데미 프로젝트의 결과물이다.

참고 문헌

[1] 안예슬, 서광규, "온라인 중고서점의 서비스 만족 요인 분석에 관한 연구", 디지털융복합연구 제11권 제11호 2013, p251 - 256.
 [2] 류용재, 조현양, 이수진, 이현우, "경기도 도서관의 연계화를 통한 활성화 방안 연구", 정책연구 연구보고서 정책연구 2009-83, 2010.02.
 [3] 이정미, "ChatGPT, 생성형AI시대 도서관의 데이터 리더십 교육에 대한 연구", 한국문헌정보학회지 제57권 제3호, 2023, p303~323.

전통문화 체험을 위한 실감형 콘텐츠 디지털 헤리티지 적용사례

우재영¹, 한진호², 위중우³, 이준원⁴, 김주언⁵, 공지훈⁶, 김장환⁷, 서채연⁸, 김영철^{9*}

^{1,2,3,4,5}홍익대학교 게임학과,

⁶툰스퀘어,

^{7,8,9*}홍익대학교 소프트웨어융합학과

{dnwodud2579¹, hanjinho216², wjw0518³}@gmail.com,

{wnsdnjs9702⁴@naver.com, juun0514⁵@gmail.com,

go400s@gmail.com⁶,

{lentoconstante⁷, chaeyun⁸, bob^{9*}}@hongik.ac.kr

Applied Practice on Digital Heritage Development for Traditional Culture Experience

Jaeyeong Woo¹, Jinho Han², Jongwoo Wee³, Junwon Lee⁴, Juun Kim⁵,

Ji Hoon Kong⁶, Janghwan Kim⁷, Chaeyun Seo⁸, R. Young Chul Kim^{9*}

^{1,2,3,4,5}School of Games, Hongik University,

⁶Toonsquare corp.,

^{7,8,9*}Dept. of Software and Communications Engineering, Hongik University

요 약

현대 사회에서는 전통문화보다는 현대화/문화화/세계화에 이슈가 크다. 그러나, 국민에 대한 전통문화의 인식을 높이려는 시도가 매우 중요한 실정이다. 젊은 세대들의 시공간적인 제약으로 전통문화의 정보 접근성이 어렵다. 우리의 유/무형 문화유산을 체험할 수 있는 대안이 필요하다. 이를 해결하고자 전통문화 체험 메타버스 모델을 제안한다. 이를 위해서, 1) 선정지역에 대한 콘텐츠 식별 및 선정, 2) 디지털 헤리티지 콘텐츠 모듈화 설계, 3) 스토리화를 위한 시나리오 정의 및 모델을 개발한다. 이는 통해 한국 전통문화 모듈화 및 재사용을 통해 쉽게 구축함으로써 전통문화 체험을 할 수 있기를 기대한다.

키워드 : 디지털 헤리티지, 실감형 콘텐츠, 무형유산, 전통문화 체험, VR, 게이미피케이션

1. 서 론

한국의 전통문화는 역사적으로 그리고 현대 사회에서도 매우 중요한 가치를 지니고 있다. 특히 한복, 윷놀이와 같은 문화적 요소들은 한국인의 정체성을 형성하는 데 큰 역할을 하며, 세대 간 전승을 통해 우리 사회의 연속성을 유지해 왔다[1]. 그러나 최근의 연구와 조사 결과들은 이러한 전통문화가 점차 소홀히 다루어지고 있으며, 특히 젊은 세대인 20대 청년층 사이에서 이러한 경향이 뚜렷하다는 점을 보여준다.

‘전통문화 진흥을 위한 정책 기초 연구 자료(2019)’에 따르면, 전체 국민 중 전통문화에 대한 관심도는 절반 수준인 50.9%에 불과하며, 특히 연령대가 낮을수록 이러한 관심도는 더욱 감소하는 경향을 보인다[2]. 이는 전통문화의 중요성에 비해 국민적 관심도가 매우 낮은 상황을 시사하며, 이러한 문화의 소중함과 가치를 재인식하고, 젊은 세대에게 전통문화를 적극적으로 알리고 확산시킬 필요가 있다. 본 논문에서는 전통문화의 현대 사회에서의 중요성이 재조명되고 있다. 특히 젊은 세대 사이에서 전통문화에 대한 인식을 증진하려는 방안을 모색하고자 무형 문화재를 실감형 콘텐츠로 제작 및 체험하고자 한다. 이를 통해 젊은 층 세대들이 전통문화를 쉽게 접할 수 있기를 기대한다. 본 논문에서 2장은 전통

문화 체험을 위한 AR 및 VR에 대해 살펴보고, 3장은 전통문화 체험을 위한 게임 메타버스 플랫폼을 언급하고, 4장은 전통문화 체험 구현, 5장은 결론을 언급한다.

2. 관련 연구

전통문화 체험을 위한 실감형 콘텐츠는 첫째, 주한 외국인들의 전통문화체험 실태와 동기 및 만족도에 관한 연구이며 한국에 거주하는 외국인들의 전통문화체험 현황을 탐구하고, 그들의 체험동기와 만족도를 분석한다. 이는 외국인들이 한국의 역사와 생활사에 친근해질 기회를 제공한다[3]. 둘째, 문화유산 체험상품의 체험속성이 문화유산 가치 인식, 지역 이미지에 미치는 영향에 관한 연구이며 문화유산이 지역 관광 활성화와 지역 브랜드 제고에 이바지할 수 있는 중요한 자원의 역할을 논의한다. 특히, 문화유산을 활용한 관광 콘텐츠가 관광객들의 참여를 유도하고 지역 문화의 인식을 높이는 방법을 제시한다[4]. 마지막으로, 전통문화 자원을 활용한 온라인 VR 체험 콘텐츠 개발 연구이며 전통문화 자원을 활용하여 VR 기술과 3D 자산을 통한 온라인 체험 콘텐츠를 개발하는 방안을 제안한다. 참여자들이 전통문화 요소와 깊이 있게 상호작용할 수 있는 인터랙티브한 경험을 창출하여 교육적 가치와 오락적 가치를 모두 높일 방안을 연구한다[5]. 이 연구들은 문화유산 활용의 현대적 접근 방식을 강조하며, VR과 같은 현대

기술을 통합하여 교육적이면서도 몰입감 있는 체험을 만드는 것이 전통문화에 대한 폭넓은 인식과 참여를 증진할 수 있다.

3. 전통문화체험 게임 메타버스 플랫폼

3.1 전통문화 표적 플랫폼에 대한 배경

전통문화에 대한 관심이 낮은 젊은 세대에게 시공간적 제약을 극복한 정보 접근성이 높은 방법으로 접근이 필요하다. 전통문화 체험 게임 메타버스 플랫폼은 가상의 세계에서 유/무형 문화유산을 한국의 전통문화를 체험할 수 있다.

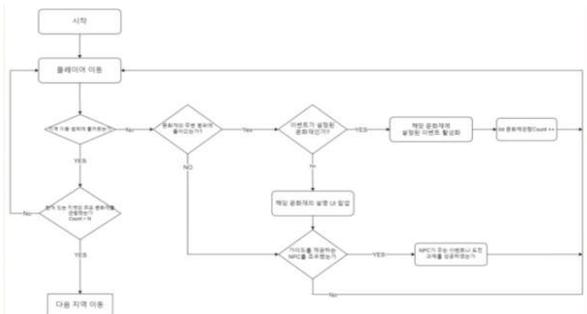
3.2 전통문화 체험을 위한 게이미피케이션

사용자가 즐길 수 있는 콘텐츠로서의 가치를 위해 게이미피케이션 요소를 추가한다. 경주의 문화유산을 선정하여 전통문화 체험을 위한 레벨 이동이다. 문화유산을 둘러보다, 팽이치기와 투호 게임이 가능하다. 그림 1은 경주 지역을 둘러보며 전통문화 체험이 가능한 종류이다. 총 3개의 전통문화 체험이 가능하다.



(그림 1) 전통문화 체험의 종류

다음은 전체 메타버스 게임 플랫폼에 대한 프로세스이다. 그림 2는 전체 메타버스 게임 플랫폼 프로세스 흐름도이다.

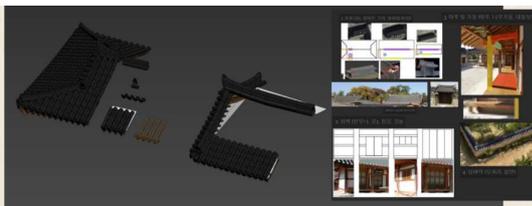


(그림 2) 전체 메타버스 게임 플랫폼 프로세스

레벨이 배치한 메인 이벤트를 진행한 후 다음 레벨로 이동할 수 있다. 메인 이벤트를 진행하기 위한 동선을 설정 후 이동 거리를 고려하여 서브 이벤트 및 게이미피케이션 요소를 배치한다.

3.3 제작 방법 모듈화

실감 나는 전통가옥을 표현하기 위해 언리얼 엔진5 한옥 3D에셋을 활용하여 모듈화한다.

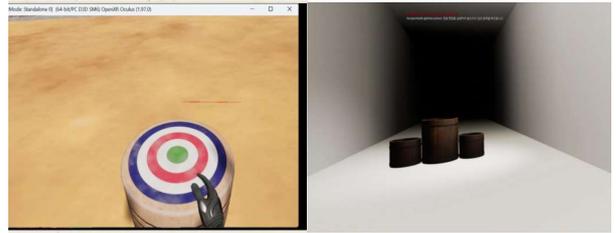


(그림 3) 한옥 3D 에셋 모듈화

언리얼 엔진5의 나나이트를 활용하여 렌더링을 효과적으로 처리하고, 루멘을 활용하여 현실적인 일루미네이션과 리플렉션을 구현하여 더 생동감 있게 표현한다[6]. 그림 3은 실감 나는 전통가옥을 표현하기 위한 한옥 3D에셋을 최적화하여 모듈화한다.

4. 전통문화 체험 VR 게임구현

4.1 팽이치기 VR 게임 및 투호 시연 VR 게임



(그림 4) 팽이치기(왼쪽)와 투호 시연(오른쪽) VR 게임

본 논문에서는 전통문화 체험을 VR 게임으로 구현한다. 그림 4는 팽이치기와 투호 시연 VR게임이다. 최 부자택으로 이동하면 팽이치기 VR 게임이 가능하고, 안압지로 이동하면 투호 VR 게임이 가능하다. 그 밖에 문화재의 모습과 얽힌 이야기, 전래동화 등을 시네마틱 장면으로 구성해 사용자에게 제공한다. 언리얼 시네마틱을 활용해 특정 이벤트를 시각적으로 표현하며 사용자가 실감형 콘텐츠에 몰입하는 것을 돕는다.

5. 결론

메타버스를 활용하여 전통문화의 특성을 살린 새로운 콘텐츠를 국내의 사용자들에게 제공함으로써, 전통문화를 보존하고 메타버스 콘텐츠 경제에 이바지한다. 이를 통해 한국 문화를 적극적으로 확산할 수 있다. 향후 메타버스 플랫폼의 특성을 활용해 전통문화 공연을 위한 가상공간을 구축하고, 오프라인 공연의 한계를 극복할 수 있다. 메타버스의 가상공간에서만 경험할 수 있는 실감형 콘텐츠는 시간과 장소의 제약 없이 편리하게 접근할 수 있으며, 실시간 상호작용과 표현의 자유성을 높인다. 이를 통해 젊은 세대가 전통문화에 대해 높은 관심을 가질 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 2024년도 문화체육관광부의 재원으로 한국콘텐츠진흥원 (과제명: 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술개발, 과제번호: RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초산산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임. 또한, 2024년 1학기 홍익대학교 메타버스 융합SW 아카데미 프로젝트의 결과물이다.

참고 문헌

- [1] 김다운, '최근 20대, 한복·전통문화에 큰 관심 없는 이유', 아이뉴스24. <http://www.inews24.com>, 2021.
- [2] 문화체육관광부. '전통 문화 진흥을 위한 정책 기초 연구 자료', 2019.
- [3] 김한샘, "주한 외국인의 전통문화체험 실태와 동기 및 만족도", 성신여자대학교 석사학위논문, 2015.
- [4] 이종욱, 박현아, 박강아, "지역 문화유산을 활용한 증강현실 기반 관광 콘텐츠 개발", 한국콘텐츠학회논문지 제18권 제6호, p122-130, 2018.
- [5] 박은비, "궁중기록화 학습을 위한 게이미피케이션 기반 가상현실 콘텐츠 개발 연구", 한국전통문화대학교 문화유산전문대학원 석사학위논문, 2023.
- [6] Epic Games, <https://www.unrealengine.com/ko/unreal-engine-5>

가상 공유 갤러리 상에서 테마와 태그 기반 효율적 이미지 검색

한진우¹, 박민혁², 홍승빈³, 우하은⁴, 공지훈⁵, 김장환⁶, 서채연⁷, 김영철^{8*}
 홍익대학교 소프트웨어융합학과
 wlsdn1232580@gmail.com¹, pnh6506@naver.com², ldb4589@naver.com³,
 whaa474@gmail.com⁴, go400s@gmail.com⁵,
 {lentoconstante⁶, chaeyun⁷, bob^{8*}}@hongik.ac.kr

Efficient Image Retrieval Method with Themes and Tags in Shared Gallery Based on Virtual Reality

Jin-woo Han¹, Min-hyuk Park², Seung-bin Hong³, Ha-eun Woo⁴, |
 Ji Hoon Kong⁵, Janghwan Kim⁶, Chaeyun Seo⁷, R. Young Chul Kim^{8*}
 Dept. of Software and Communications Engineering, Hongik University

요약

기존의 웹사이트의 정보 분류 방식은 다양한 주제와 태그를 통해 정보를 체계적으로 분류하는 방식이다. 현재 가상 환경 내에서의 사진 갤러리에서 효율적 탐색 방법이 필요하다. 이 문제를 해결하기 위해서 가상 환경 내에서의 주제 및 태그 방식 적용을 제안한다. 이를 통해 사용자는 사진을 주제별로 구분하고, 세부적으로는 태그를 통해 분류하여 더욱 효율적으로 가상 환경 내에서 이미지를 검색할 수 있다. 이러한 접근 방식은 단순히 이미지를 찾는 것을 넘어서, 사용자가 새로운 발견과 경험으로 사용자에게 더 직관적이고 즐거운 사진 탐색 경험을 제공하며, 갤러리 사용의 만족도를 높일 수 있다.

1. 서론

일상생활에서 다양한 웹사이트를 이용하면서 특정 주제를 기준으로 분류된 내용을 쉽게 찾아볼 수 있음을 알 수 있다. 또한, 주제 내에서도 태그를 이용해 세부적으로 분류되어, 우리가 찾고자 하는 정보나 비슷한 내용을 더욱 편리하게 찾아볼 수 있다. 예를 들어 네이버의 경우 업데이트 날짜를 주제로 분류하고 무신사는 옷의 종류 그 속에서 스타일로 나눈다 [1]. 유튜브 또한 음악, 영화 등 영상의 종류를 주제로 분류한다 [2]. 우리는 이러한 방식을 갤러리에 적용한다. 사진의 주제와 태그를 통한 분류를 통해, 우리가 원하는 이미지를 더욱 손쉽게 찾아볼 수 있게 한다. 이를 통해, 사용자는 원하는 이미지를 찾는 동시에 새로운 발견과 경험을 쌓을 수 있다. 기존에 존재하던 애플리케이션 제품들의 서비스 방식들을 참고하고 사용자들이 불편함을 느끼는 문제점들을 개선하고, 주제 기능으로 기존의 제품들과 다른 차별적인 UI/UX 디자인을 제공하는 웹 애플리케이션을 설계 및 구현하여 차별화된 서비스를 제공하고자 한다. 2장에서는 관련 연구에 대해 언급하고, 3장에서는 제안한 효율적 이미지 검색 갤러리 구현 방법에 대해 언급하고, 4장은 결론을 언급한다.

2. 관련 연구

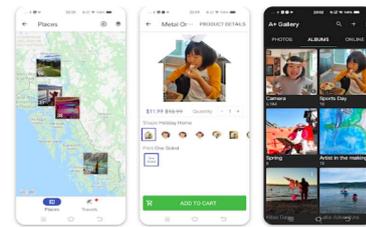
2.1 모두의 갤러리

모두의 갤러리는 간편한 사진 앨범 앱이다. 일반 갤러리와 공유 갤러리를 구분하여 생성한다. 초대하기에서 다양

한 방식으로 공유 갤러리에 초대한다. 사진이나 동영상을 보고 댓글을 남긴다 [3].

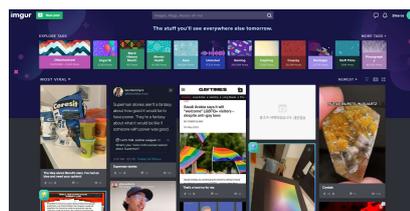
2.2 A+ 갤러리

A+ Gallery는 Android 휴대전화를 위한 사진 갤러리 앱이다. 이 애플리케이션에서는 HD 사진 보기, 사진 검색 및 앨범 관리를 빠르게 실행한다. 그림 1은 A+ 갤러리의 예시이다 [4].



(그림 1) A+갤러리

2.3 Imgur

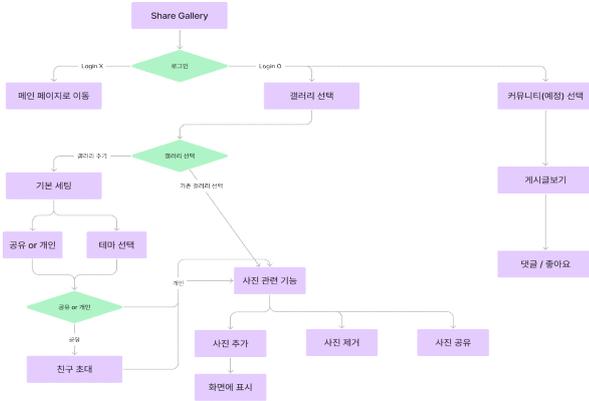


(그림 2) imgur 이미지공유 커뮤니티

Imgur사이트는 미국의 이미지 공유 커뮤니티이다. 사진

뿐만 아니라 동영상까지 올려 유튜브 쇼츠, 틱톡과 같은 다양한 콘텐츠가 있다. 그림 2는 Imgur 이미지 공유 커뮤니티의 화면이다 [5]. 차별적인 기능을 활용하기 위해 관련된 다양한 기능들을 적용한다.

3. 효율적 이미지 검색 방법

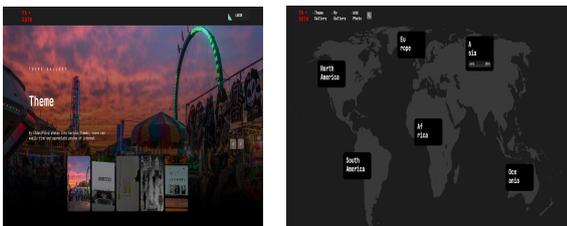


(그림 3) 가상 공유 갤러리 주제/태그 기반 이미지 검색 흐름도

그림 3은 가상 공유 갤러리 상에서 주제와 태그 기반으로 효율적 이미지를 검색할 수 있는 전체 흐름도이다.

사용자가 원하는 태그로 갤러리 내 사진을 분류할 수 있는 기능을 기반으로, 해당 기능에 대한 접근성을 높이는 UI로 갤러리를 구현한다. 우리가 제안한 효율적 이미지 검색 방안은 주제 기능으로 기존의 제품들과 다른 차별적인 UI/UX 디자인을 제공하는 웹 애플리케이션을 설계 및 구현하여 사용자들에게 새로운 경험을 제공하고자 한다 [6].

4. 효율적 이미지 검색 구현



(그림 4) 메인페이지(왼쪽)와 주제 페이지(오른쪽)

효율적 이미지 검색을 위해 구현된 웹페이지를 살펴본다. 메인 페이지는 사이트의 메인 기능들이 나열되어 사용자가 사이트에서 제공하는 다양한 기능들을 쉽게 파악하고, 필요한 정보를 신속하게 찾을 수 있다. 그림 4에서 왼쪽은 가상 갤러리 웹의 메인 페이지이고, 오른쪽은 주제 페이지이다. 주제 페이지는 갤러리에 업로드된 이미지 목록들을 주제 태그별로 확인할 수 있는 페이지이다. 주제 페이지에서는 원하는 주제로 갤러리를 구성하고 사진을 분류 배치하여 각 주제에 어울리는 UI를 제공해 사용자에게 다양한 경험을 제공한다. 가상 웹 갤러리의 VR 사진 갤러리 기능이다. 가상 현실 속 갤러리 공간을 구축하여 실제 갤러리에 온 것과 같은 느낌을 받을 수 있는 가상공간이다. 그림 6은 VR 사진 갤러리이다. 이 가상 갤러리는 현실과

가상이 만나는 예술 공간이다. VR기기를 활용한 VR 사진 갤러리 구축은 사용자에게 실제 갤러리에 온 것 같은 몰입감을 제공하며, 현실감 있는 환경을 통해 공간 제약 없이 다양한 작품을 전시할 수 있게 한다.



(그림 5) VR 사진 갤러리

또한, 작가와 관람객 간의 소통을 하여, 가상 현실 속에서도 실제 미술관처럼 풍부한 예술적 경험과 상호작용을 제공한다 [7].

4. 결론

가상 공유 갤러리에서 다양한 주제로 분류된 사진들을 사용자 지정 태그 기능을 통해 빠르고 쉽게 검색하고 분류할 수 있는 효율적인 이미지 검색 방법을 제안한다. 이 방법을 통해 다양한 주제로 분류된 사진들을 쉽게 검색하고 분류할 수 있다. 사용자는 특정 주제의 사진을 통해 새로운 정보를 얻거나 공유할 수 있다. 또한, 시공간에 구애받지 않고 공통의 커뮤니티를 형성하여 다양한 볼거리와 즐거움을 누릴 수 있기를 기대한다.

ACKNOWLEDGEMENT

본 연구는 2024년도 문화체육 관광부의재원으로 한국콘텐츠진흥원(과제명: 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술 개발, 과제번호: RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임. 또한, 2024년 1학기 홍익대학교 메타버스 융합SW 아카데미 프로젝트의 결과물 및 홍익대학교 소프트웨어 융합학과 학부생들의 종합설계 프로젝트 결과물이다.

참고 문헌

- [1] 박소연, “주요 포털들의 서비스 분류체계 비교 분석”, 한국문헌정보학회지 학술저널 제44권 제2호, 2010. p241-262.
- [2] 김용우, 전성배, 정유철, “유튜브 메타정보를 이용한 자동 주제 분류 고찰”, 한국컴퓨터정보학회 학술발표논문집 학술대회자료, 제29권 제1호, 2021. p349-351.
- [3] 모두의 갤러리, <https://modoogall.com>
- [4] A+ 갤러리, 구글 플레이스토어.
- [5] imgur, <https://imgur.com/>
- [6] 홍순재, “UI/UX 디자인 교육 활용에 관한 연구: 피그마를 중심으로”, 한국디자인리서치학회 학술저널 제8권 제2호, 2023. p34-43.
- [7] 변민주, “VR/AR을 기반으로한 수원시립미술관 디지털워킹 아트의 공간체험형 융합스토리텔링 방법론 연구”, 한국과학예술융합학회 학술저널, Vol.39 No.5, 2021. p119-131.

전기차량에 적용된 케블라 섬유 열선 히터 불량 검출기 개발 및 제작

¹ 최우성, ² 박재완, ¹ 이준희, ¹ 성명훈, ¹ 김재원, ¹ 조소현, ¹ 우 주, ¹ 정재훈
¹ 군산대학교 소프트웨어학부, ² (주)재현이노텍
 cdt0226@naver.com, jhinnotek@naver.com, zmdzmd0541@naver.com,
 choshin5016@daum.net, kjw8541@naver.com, shjo960@naver.com,
 oojoo@kunsan.ac.kr, jh7129@kunsan.ac.kr

Development and fabrication of a fault detector for Kevlar fiber heating elements applied in electric vehicles

Choe Useong, Pak Jaewan, Lee Junhui, Jo Sohyen,
 Woo Ju, Jeong Jaehun, Seong Myeonghun, Kim Jaewon
 Kunsan University Department of Software, Jaehyun Innotech Co., Ltd.

요 약

본 연구에서는 날씨 및 외부요인에 의해 전기자동차 자체 내부 온도가 높고 열선에 의한 화재, 화재로 이어질 수 있는 가능성을 배제하고자 케블라 섬유 열선 히터 검출기를 개발 및 제작하여 저항값을 측정함으로써 박음질시 발생하는 문제점을 발견하고 방지할 수 있다.

1. 서 론

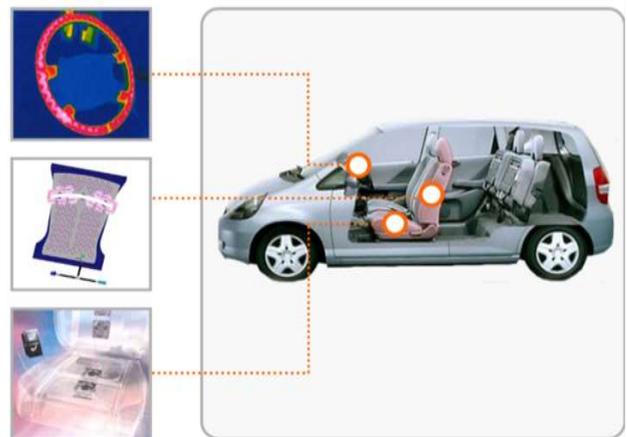
전기자동차의 사용자가 점차 늘어가고 있으면서 화재가 발생했을 경우 쉽게 진압하지 못해 골머리를 앓고 있는 문제점을 보완하기 위해 다양한 연구를 시도하고 있다. 대표적으로 화재의 원인인 배터리는 자체 발열로 배터리 팩 내부 온도는 상당히 높으며 이를 방열하기 위한 냉각 장치를 꾸준히 연구 개발하고 있다. 그럼에도 불구하고 배터리로 인한 차량 화재는 막지 못하는 편이며, 여름 같은 외부 온도가 높은 환경과 차량의 배터리로 인한 내부의 높은 온도까지 더해진다면 차량 자체적인 온도는 상당히 높은 온도이다. 여름과 차량 주행 중 배터리 온도 등 2가지 만으로도 높은 온도까지 올라가며, 추가적으로 각종 조향 장치에 히팅 시스템이 장착되어 부위별 히팅이 되는 발열 열선이 존재한다. 히터 열선이 동작하게 되면 자체만으로도 60도 이상의 온도가 발열되며 이때 부위별 열선의 내구성에 문제가 있다면 열화로 인한 합선으로 이어지게 된다. 즉 전기자동차는 열에 의한 열손 및 화재 예방에 관련된 모든 부품으로 구성되어 있어야 하고 본 과제를 통해 케블라 히팅 열선 불량 검출기를 개발하여 2차 가공 시 가죽 감싸기 후 발생하는 불량 유형을 검출하고자 한다.

2. 본론

2.1 개발의 필요성

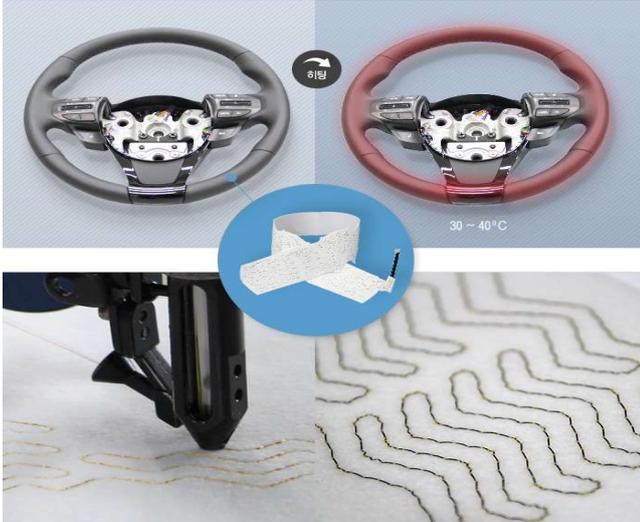
케블라 브랜드 섬유는 고강도와 경량성을 겸비한 듀폰의 혁신적인 기술로서, 다양한 소비자 및 산업용 제품의 성능을 향상시켜주고 미국 듀폰사는 1971년에 케블라를 출시하여 파라아라미드 계열의 슈퍼 섬유 시대를 열었다. 케블라의 초창기 생산량은 1971년에 200t/year, 1974년에 2,700t/year 수준으로 70년대에는 그 용도 개발에 있어 우주항공 소재, 타이어 코드,

방호복 등의 용도로 개척했다. 전기자동차는 배터리 발열과 외부 온도의 영향으로 차량 자체의 온도가 매우 높고, 차량 내부의 조향 장치에 히팅 시스템이 장착되어 있어 부위별 히팅이 되는 발열 열선이 존재한다. 히팅 열선 동작시 60도 이상의 높은 온도로 발열하게 되고, 내구성에 문제가 있다면 열화로 인한 합선으로 이어지게 된다. 주행 중인 전기자동차의 내부는 배터리 열과 외부 열 등으로 인해 온도가 높고 이때, 어느 부분에서 합선이 발생한다면 화재로 이어질 수 있으며, 전기자동차의 화재는 진압하기 매우 어렵다. 또한 운전자가 직접 손으로 잡는 부위이기에 열로 인한 화재 주의가 필요하며 온도 컨트롤러 불량이면 온도 제어가 되지 않아 온도가 계속 올라가 소손된다. 따라서 전기자동차 안전성을 향상시키기 위해 케블라 섬유 열선 히터 불량 검출기를 개발하고자 하고 스티어링 휠에 히팅 열선을 탑재할 때 발생하는 불량 유형을 검출하고자 한다 [1-5].



(그림 1) 전기자동차 내부 히팅 시스템 위치

2.2 케블라 섬유 히터 열선 테스트 대상 선정

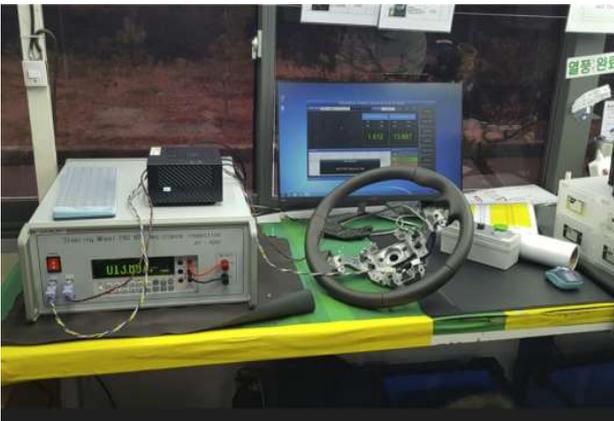


(그림 2) 스티어링 휠

전기 자동차 내부에선 히팅을 사용하는 부품 중 가장 취약한 부품인 스티어링 휠을 대상으로 검사 진행할 계획이며 차량 내부 태양빛에 가장 많은 노출이 되는 부분이 핸들 부분이며, 열을 많이 받은 상태에서 히팅이 이루어지기 때문에 품질 관리에 매우 중요하고 운전자가 직접 손으로 잡는 부위이므로 주의가 필요하다.

3. 실험

실제로 현장에서 단품으로 제작 후 검사한 결과로 단품으로 검사할 시 불량률은 3% 내외이며 단품 생산 후 큰 문제점이 발생되지 않는 것을 확인했다. 문제점은 단품이 아닌 단품 열선 라인에 가죽을 입히고 모양대로 박음질 후 케블라 섬유에서 벗겨진 구리 선에 대한 불량이 발생하는데 2차 가공 시 발생하는 불량 유형 검출 시스템이며, 저항을 측정하여 불량을 검출한다. 케블라 섬유에 구리 라인이 직병렬로 구성되어 있어 검출이 불가능하다. 따라서 케블라 섬유를 제외한 구리 라인 직병렬 조직에서의 저항 검출 시스템을 개발했다.



(그림 3) 케블라 섬유 열선 히터 저항 측정

4. 결론

본 연구에서는 열선 파손으로 인한 차량 화재를 예방하기 위해 케블라 섬유 열선 히터 불량 검출기를 개발하였다. 단품 열선을 스티어링 휠에 감싸고 가죽을 입히는 과

정에서 불량이 발생하는 것을 확인하였다. 따라서 2차 가공이 끝난 스티어링 휠의 케블라 열선 불량을 검출한다. 열선에 대한 저항과 케블라 섬유 열선 히터에 탑재되어 있는 NTC 센서 저항을 측정하여 양품과 불량품을 검출할 수 있다. 제네시스 G90 차량의 스티어링 휠을 이용하여 실험을 진행하였으며, 90% 이상의 검출 결과를 확인하였다. 케블라 열선 히터 불량 검출기를 통해 스티어링 휠 오류로 인한 리콜과 열선 불량으로 인한 차량 화재 사고를 예방할 수 있을 것으로 기대된다.

참고 문헌

- [1] Singh, Thingujam Jackson, and Sutanu Samanta. "Characterization of Kevlar fiber and its composites: a review." *Materials Today: Proceedings* 2.4-5 (2015): 1381-1387.
- [2] Chatzi, E. G., and J. L. Koenig. "Morphology and structure of Kevlar fibers: a review." *Polymer-Plastics Technology and Engineering* 26.3-4 (1987): 229-270.
- [3] Saw, Lip Huat, Yonghuang Ye, and Andrew AO Tay. "Integration issues of lithium-ion battery into electric vehicles battery pack." *Journal of Cleaner Production* 113 (2016): 1032-1045.
- [4] Chinmasamy, Vivekanandhan, et al. "Characterization on thermal properties of glass fiber and kevlar fiber with modified epoxy hybrid composites." *Journal of materials research and technology* 9.3 (2020): 3158-3167.
- [5] 김창수, 전형훈, 김명관, and 정용식. "리튬이온 (폴리머) 배터리의 화재위험성 연구." *한국화재감식학회 학회지* 11, no. 3 (2020): 19-27.

사사 : 이 성과물은 2024년도 정부(과학기술정보통신부)의 재원으로 연구개발특구진흥재단-2024년 강소특구 지역특성화 육성사업(전북 군산특구)사업 의 지원을 받아 수행된 연구임(2024-IT-RD-0019-01-101)

애장품 소유 감성 메커니즘을 적용한 게임 제작 사례

손희완, 은광하

상명대학교 디지털콘텐츠전공

e-mail : marry1315@naver.com, kheun0310@smu.ac.kr

Game Development Applying the Emotional Mechanisms of Ownership

Hee-Wan Son, Kwang-Ha Eun

Major in Digital Contents, Sangmyung University

요약

본 연구에서는 사회적 인간의 성장과 관련하여 애장품의 소유 감성에 착안하여 메커니즘을 설계하여 플레이어 몰입도를 높일 수 있는 디자인 접근이 적용된 게임 제작 사례이다. 게임 주인공의 잃어버린 마음을 간헐적인 감옥에 대입하여 감옥을 탈출하는 모험이라는 게임 스토리라인을 기반으로 대표적인 애장품인 유년시절의 장난감을 게임 플레이상에서 주요한 상호작용 요소로서 적용하고 제작된 게임이 소유 감성의 효과성을 제시했는지 검증해 보고자 한다.

1. 서론

게임산업의 지속적인 발전과 함께 최근에는 소규모의 개발팀에 의한 인디게임 개발이 확장되고 있으며 인디게임의 주요 접근 방식은 게임 몰입성의 핵심인 메커니즘의 참신성이다.

본 논문은 애장품 소유 감성을 게임 디자인에 적용한 제작 사례로서 애장품이 가지는 감성적 가치는 사용자에게 특별한 의미와 가치를 부여하며, 이러한 감성을 게임 메커니즘에 적용함으로써 플레이어의 몰입도를 높일 수 있는 디자인 접근이라고 할 수 있다. 본 연구에서는 애장품 소유 감성의 게임 메커니즘을 적용하기 위하여 장난감 오브젝트인 게임 요소로 적용하여 획득과 스킬 사용의 메커니즘을 설계하여 프로토타입을 제작하여 이용자를 대상으로 플레이한 후, 소유 감성의 메커니즘이 게임의 몰입성에 효율적으로 반영되는지 검증해 보고자 한다.

2. 본론

2.1 관련 게임의 메커니즘 적용 사례 분석

게임 제작에 들어가기에 앞서, 레퍼런스 게임들을 분석해 보았다. 첫 번째는 팀 체리의 ‘할로우나이트’이다. 팀 체리는 2인으로 구성된 호주의 인디게임 개발팀이다. 벨레들의 왕국인 할로우네스트를 배경으로, 망해버린 신성동지를 기억을 잃은 주인공이 탐험하면서 수수께끼로 가득한 세계를 알아가며 자신의 소임을 알아낸다. 끝에서는 주인공은 사실 신성동지를 멸망으로 이끈 광휘를 봉인하는 그릇인 걸 알고, 그를 봉인하는 스토리이다. 이 작품의 특징적인 것은 체력을 전부 잃게 되면 가면이 깨지고 재화를 전부 잃어버린다는 점이다. 이 점은 게임 초반에 치명적으로 다가오는데, 아무것도 모르는 세계에 떨어져 지도를 수집해야 하고, 강한 몬스터를 상대해야 해서 재화가 필요하

다. 이 점은 플레이어에게 스트레스를 주면서도, 죽지 않으려고 하는 스틸감을 준다. 두 번째는 ‘레밍즈(Lemings)’이다. 캐릭터는 계속해서 좌에서 우로 걷기만 한다. 그러다 벽 같은 것에 닿으면 반대 방향으로 걸어가게 된다. 뽀족한 가시 몇 함정들을 게임을 하는 유저는 하단에 특수한 능력을 사용하는 레밍을 소환하여 장애물이나 지형을 극복하여 레밍을 제한 시간 내에 일정 수를 다음 스테이지로 보내야 한다. 여기서 유저는 타이밍에 맞게 적절한 능력을 갖춘 레밍을 소환해야 한다.

2.2 게임 시스템 기획

1) 게임 기본 시스템

레퍼런스 게임 분석을 바탕으로 게임 시스템을 설계하였다. 게임의 기본적인 플레이 방법은 기존의 ‘마리오’, ‘캐슬 배니아 시리즈’ 등으로 사용자에게 익숙한 횡 방향 플랫폼 게임으로 기획하였다. 게임의 스토리는 가족들의 불화로 생긴 트라우마로 자신만의 세계에 빠져 사는 주인공이 장난감 세계로 빨려 들어가게 된다. 감옥에서 탈출하는 모험을 통해 스스로 마음을 열고 희망을 되찾는 과정을 플레이어가 느낄 수 있게 할 예정이다. 이 게임 플레이의 핵심 요소는 주인공이 장난감을 소환하면서 협력하는 것을 통해 자신에게 처한 여러 가지 문제들을 하나하나 해결해 나가는 점이다.

2) 소유감성 메커니즘의 장난감 시스템

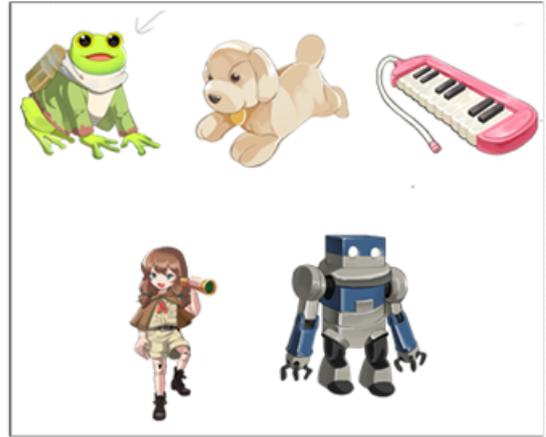
캐릭터와 함께하는 장난감의 협력 플레이가 게임의 핵심 메커니즘이라고 할 수 있다. 예를 들면 캐릭터 앞을 가로막은 바위 블록과 그 너머에 있는 몬스터를 사냥해야 하는 상황을 가정해 보자. 우선 캐릭터는 키보드 e를 눌러 장난감 목록에서 장난감을 선택할 수 있다. 바위 블록을 파괴할 수 있는 개구리 캐릭터를 소환한 후, 개구리 장난

감에는 3가지 능력을 쓸 수 있고, 1번 2번 3번 스킬 칸이 있다. 3번 스킬을 활용하여 폭탄을 소환하여, 바위 블록을 부순다. 플랫폼을 올라가면, 봉제 인형을 한 몬스터를 마주하게 된다. 봉제 인형을 없앨 수 있는 강아지 인형으로 장난감을 전환하여 강아지 장난감을 소환해 적을 제거한다.

(표 1) 장난감 스킬 목록

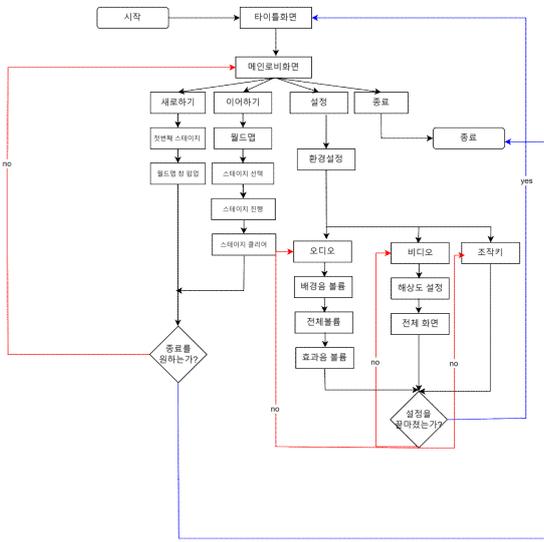
장난감 이름	플라스틱 강아지 장난감	메탈 로봇 장난감	닌자 개구리 봉제인형	휴대용 악기	탐험가 인형
능력1	강아지의 후각	방패	개구리 점프	혼란	지도
능력2	이속 증가	로봇팔	개구리 혀	매혹	햇불
능력3	물기	레이저	폭탄	식물 성장	뱃줄

왼쪽에서부터 닌자 개구리 봉제인형, 플라스틱 강아지 장난감, 휴대용악기, 탐험가 인형, 메탈로봇장난감이다.



(그림 2) 장난감 캐릭터 일러스트

3) 게임 플로우차트



(그림 1) 게임 <잃어버린 마음 > 플로우 차트

게임을 실행하게 되면 타이틀 화면을 메인 로비 화면이 나온다. 만약 게임을 처음 실행하게 되면, 새로 하기 설정 종료 창이 메인화면에 뜨게 되고, 이전에 플레이했다면 이어 하기 메뉴도 화면에 팝업된다. 게임을 새로 시작하면 첫 번째 스테이지가 나오고 첫 번째 스테이지에서 게임의 기본적인 조작법과 주인공이 어떻게 장난감 세계에 빠지게 되었는지에 대한 스토리에 대해 알게 된다. 첫 번째 스테이지를 클리어하면 다음 스테이지로 이어지는 팝업이 뜬다. 이어하기로 시작한 플레이어는 마지막으로 실행 종료했었던 스테이지에서 시작해 스테이지를 진행하고 클리어하면서 게임이 진행된다. 스테이지 끝부분으로 가면 다음 스테이지로 갈지 종료할지에 대한 창이 팝업되고, 종료를 하면 게임이 종료된다. 설정 창은 오디오 볼륨, 해상도 설정, 조작키에 대한 창으로 구성되어 있다. 종료 창을 클릭하면 종료한다.

4) 게임 제작 리소스 및 구현 화면

맨 왼쪽은 플레이어가 선택한 장난감 캐릭터이고, 그 다음 3개의 버튼들은 그 장난감이 가지고 있는 능력이다.



(그림 3) <잃어버린 마음> 게임 플레이 화면

4. 결론

플레이어 캐릭터의 몰입성을 위해서 소유 게임 메커니즘을 상호작용 요소로 적용하여 게임의 몰입성이 확대되는지 확인할 필요성이 있다. 따라서 향후 사용자 체험을 통해서 제시한 소유 감성 메커니즘의 효과성을 검증해 보고자 한다.

참고 문헌

[1] 안진경(2016.02.12.) 로그라이크 게임에 나타난 영속적 죽음의 매커니즘 연구, 한국게임학회 논문지 제 16권 제 1호, 33-42
 [2] 이정엽(2014.08) 인디게임의 인터랙션 매커니즘과 서술 방식 연구, 한국 게임학회 논문지 제 14권 제 4호 17-26
 [3] 윤찬, 우탁 (2018.12.06.) 게임 과몰입과 중독에 영향을 미치는 게임 요소와 매커니즘 연구, 한국 게임학회 논문지 제 18권 제 6호,131-142

인터랙티브 드라마 <Deleter> 개발사례 분석

김도혜, 조성주, 교신저자 허영
 상명대학교 디지털콘텐츠전공
 e-mail : kd0.ohye@gmail.com

Analysis of Interactive drama <Deleter> development case

Kim Do hye, Jo seong ju, Hur young
 Department of Desital Contants at Sangmyung University

요 약

인터랙티브 드라마 <Deleter>는 시청자의 선택을 통해 스토리 전개에 참여하도록 하여 몰입감을 높이고, 짧은 러닝타임으로 가볍게 즐길 수 있는 목적으로 제작되었다. 본 콘텐츠는 실험적인 영상, 모션그래픽, 선택지로 이야기와 다양한 시각적 경험을 제공하고 이야기의 분기점, 그에 따른 시청자는 다양한 콘텐츠를 자연스럽게 융합하여 즐길 수 있도록 설계하였다. 이를 통해 시청자에게 개인화된 경험을 제공하고, 기존 인터랙티브드라마와 차별화된 독특한 경험을 제공하고자 하였다.

1. 서 론

디지털 기술의 발전은 콘텐츠의 종류를 발전시켰고, 그의 발전에 따라 TV드라마, 애니메이션, 영화, OTT 등의 다양한 매체와 스토리텔링 양식이 탄생하였다. 이렇게 탄생한 다양한 양식들은 상호 간의 경계를 넘나들며 콘텐츠 간의 융합을 이루기도 한다. VFX의 발달로 인한 영화같은 게임의 등장과 OTT서비스의 대중화로 인한 게임과 같은 상호작용성을 가진 영화 콘텐츠들의 등장이 대표적인 예시이다.

앞서 설명하였듯이 기술의 발달로 인해 제작자와 소비자 사이의 쌍방향적 소통이 가능해짐에 따라 미디어 콘텐츠에서 인터랙티브 스토리텔링 양식을 많이 활용하려는 추세이다.

본 논문은 소비자의 흥미와 몰입을 이끌 수 있는 인터랙티브 서비스를 통해 어떻게 소비자에게 더욱 다양한 시각적 경험을 제공할 수 있을지에 목적이 있다.

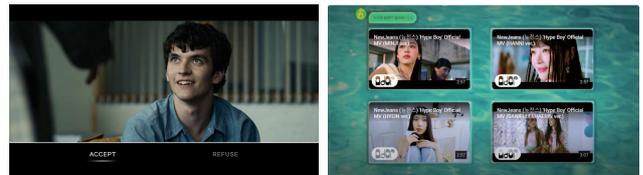
2. 인터랙티브 드라마 사례분석

2.1 인터랙티브 콘텐츠

인터랙티브 콘텐츠는 인터랙티브 미디어 기술을 기반으로 제작된 모든 종류의 쌍방향적 콘텐츠를 말한다. 사용자와의 상호작용을 기반으로 정보가 전달되며, 사용자의 선택이나 입력에 따라 결과나 경험이 달라지는 콘텐츠를 의미한다. 이는 단순한 정보 전달을 넘어 소비자가 콘텐츠에 몰입하고 적극적으로 참여하게 만드는 특징을 가진다.

인터랙티브 콘텐츠는 다양한 형태로 제공되며 대표적으로 게임, 가상현실 및 증강현실, 인터랙티브 비디오 등이 있다. 넷플릭스의 <블랙미러: 밴더스내치>와 뉴진스의 <하임보이> 뮤직비디오가 인터랙티브 비디오의 대표적인 예시이다. <블랙미러: 밴더스내치>는 2018년에 공개된 인터랙티브 영화로 인간의 자유의지와 기술의 영향을 주제로 삼고 있다. 시청자가 선택하는 옵션에 따라 이야기의

흐름과 결말이 달라지고, 시청자가 어떻게 선택하느냐에 따라 다섯 종류의 엔딩을 보여준다.



(그림 1) <블랙미러: 밴더스내치>, (그림2) 뉴진스 <하임보이> 뮤직비디오

뉴진스는 <하임보이> 뮤직비디오를 인터랙티브 형식으로 공개했다. 처음 뮤직비디오의 인트로가 끝나면 각 멤버가 주인공인 4개의 뮤직비디오가 선택지로 주어진다. 시청자들은 원하는 순서대로 시청하면서 뮤직비디오의 전체 서사를 파악하고 즐길 수 있다.

2.2 인터랙티브 콘텐츠의 이점과 한계

인터랙티브 콘텐츠는 소비자, 즉 시청자와 제작자의 경계가 흐릿해진다는 특징을 가진다. 이를 통해 사용자에게 맞춤형 경험을 제공하며, 그들의 선택과 행동이 콘텐츠 결과에 영향을 미치며 직접 콘텐츠와 상호작용하므로 단순히 소비하는 콘텐츠보다 몰입도가 높다. 이처럼 시청자가 콘텐츠에 능동적으로 참여하게 만들기 때문에 더 오래 콘텐츠에 머무르게 하고 브랜드나 플랫폼에 대한 충성도를 강화할 수 있다. 또한, 실시간 피드백을 통해 즉각적인 학습 효과를 제공하고 마케팅 및 교육 분야에서 효과적으로 활용될 수 있다.

인터랙티브 콘텐츠는 많은 이점을 가지고 있는 스토리텔링 형식이지만 기술적인 제약이 있을 수 있고, 콘텐츠의 서사의 깊이가 얕아질 수 있다. 높은 상호작용성을 제공하기 위해서는 특정 디지털 기술이 요구되며, 모든 시청자나 디바이스에서 원활히 작동하지 않을 수 있다. 인터랙티브 콘텐츠는 여러 가지 선택지를 제공해야 하기에, 각각의 이야기나 결말이 충분히 깊이 있게 다뤄지지 못할 가능

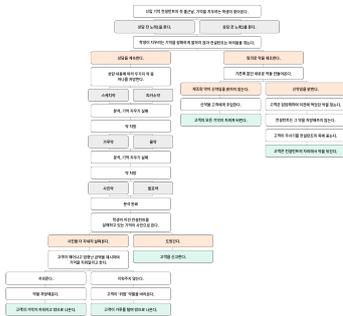
성이 높다. 하지만 이런 제한된 서사를 시청자의 선택에 따른 모든 경로와 결과를 준비해야 하기에 콘텐츠 개발이 복잡하고 타 콘텐츠에 비해 비용이 많이 든다.

결국 인터랙티브 콘텐츠는 시청자의 흥미를 유발하고 몰입감을 높여 개인화된 경험을 제공할 수 있는 강력한 스토리텔링 형식이지만, 제작 비용과 사용자 피로도 등의 한계를 극복해야 하기도 한다.

3. 인터랙티브 드라마 <Deleter> 기획

3.1 인터랙티브 드라마 <Deleter>의 특징

시청자가 단순히 수동적으로 콘텐츠를 감상하는 것이 아니라, 스토리의 전개에 참여하고 있다는 느낌을 제공하여 몰입도를 높여줄 것이다. 이때 계속해서 시청자에게 선택을 요구하면 일부 시청자에게 피로감을 줄 수 있기에 콘텐츠의 시작부터 엔딩까지의 길이를 최대한 짧게 설정하여 시청자들이 가볍게 즐길 수 있는 스낵 콘텐츠로 제작한다.



(그림 3) <Deleter>의 플로우차트

또한 개인화된 경험을 제공하여 시청자들의 상상력을 자극하려 한다. 선택지를 기점으로 이야기가 나누어진다는 점을 활용하여 하나의 미디어 콘텐츠가 아닌 다양한 형태의 미디어 콘텐츠를 함유하도록 유도한다. 이처럼 하나의 콘텐츠에서 다양한 시각적 경험을 이끌어내는 것이 타 인터랙티브 드라마와의 차별점이자 목표이다.



(그림 4) 인터랙티브 드라마 <Deleter> 장면

3.2 <Deleter> 속 콘텐츠

인터랙티브 드라마 <Deleter>에서 제공하는 콘텐츠는 크게 실사 촬영물과 모션그래픽으로 나눌 수 있다. 이야기 진행에 기반이 되는 실사 촬영은 두 명의 배우가 대화를 주고받으며 시청자로 하여금 스토리의 전체적인 세계관과 흐름, 그리고 인물의 숨겨진 감정을 묘사한다. 인터랙티브 드라마라는 실험적인 구조에 이어서 드라마 미술포도 실험적으로 구성하였다. 인물과 배경에 명도 대비를 주고, 장소의 배경을 흰색으로만 배치하여 인물들에게 시선이 가도록 유도하고 실험적인 특성을 강조한다.



(그림 5) <Deleter> 실사 장면

총 6개의 모션그래픽을 제작하여 실사 촬영물과 함께 구성한다. 학생의 기억을 표현하는 도구로 사용하였으며 학생에게 주입되는 약의 강도에 따라 달라지는 모션그래픽을 시청할 수 있다. 각 모션그래픽은 세 종류로 나뉘는데 스케치와 도형을 기반으로 제작된 모션그래픽, 애니메이션을 기반으로 제작된 모션그래픽, 그리고 실사 촬영물을 기반으로 제작된 모션그래픽으로 나뉜다. 시청자의 선택에 따라 각 종류의 모션그래픽을 선택하여 시청할 수 있다.



(그림 6) <Deleter> 도형을 기반으로 한 모션그래픽 장면



(그림 7) <Deleter> 애니메이션을 기반으로 한 모션그래픽 장면



(그림 8) <Deleter> 실사 촬영물을 기반으로 한 모션그래픽 장면

4. 결론

본 인터랙티브 드라마는 일반적으로 인터랙티브 콘텐츠에서 찾아볼 수 있는 시청자의 몰입도 증가와 개인화된 경험 제공을 포함하여, 보통의 인터랙티브 드라마의 길이에 비해 짧게 구성되어 시청자들이 콘텐츠에 대해 피로를 느끼지 않고 콘텐츠를 즐길 수 있다. 또한 선택지를 기점으로 분기점이 생긴다는 점을 이용해 하나의 콘텐츠 안에서 다양한 형태의 콘텐츠를 융합이 자연스럽게 진행되고, 이를 통한 시청자의 다양한 시각적 경험을 기대한다.

참고문헌

[1] 남승희. “실사 촬영물에 기반한 인터랙티브 드라마 연구”, 『디지털스토리텔링연구』 제3권 제1호, 2008.216-239

YOLOv8기반 배전공가 관리 시스템

이훈제*, 조성민*, 정세훈**, 심춘보*

*국립 순천대학교 인공지능공학부

**국립 순천대학교 컴퓨터공학과

e-mail : leehj4087@naver.com, cbsim@scnu.ac.kr

YOLOv8-based Distribution Substation Management System

Hun-jea Lee*, Seong-Min Jo*, Se-Hoon Jung**, Chun-Bo Sim*

*Dept. of Artificial Intelligence Eng., Suncheon National University

**Dept. of Computer Eng., Suncheon National University

요약

본 논문은 전봇대에 무단으로 설치된 전선을 탐지하기 위한 연구이다. 전선 객체는 얇고 복잡한 형태로 되어 있어 탐지가 잘되지 않는 어려움이 있다. 이 연구에서는 YOLOv8l-seg 모델을 기반으로 객체를 탐지하고 2DCNN 알고리즘으로 전선 객체의 변화를 탐지한다. 데이터 품질 개선을 위해 Polygon 기법과 데이터 증강을 했다. YOLOv8l-seg의 Anchor-free를 조정하여 다양한 전선을 탐지하고 2DCNN의 유클리드 거리 차이를 계산하여 두 이미지의 특징을 비교해 전선의 변화 여부를 판별한다. 실험 결과 개별 전선을 정확하게 구분하고, 두 시점의 이미지 변화를 식별할 수 있었다.

1. 서론

최근 3년간 한국전력은 43조 원의 누적 적자를 기록하였고, 전국에 설치된 약 4만 km의 무단 통신선을 정비할 계획을 발표했다[1]. 배전공가는 전력 공급에 필수적인 인프라로, 수작업에 의존한 정기 점검은 많은 인력과 시간이 소요되며, 미세한 변화를 감지하기 어렵다. 따라서 효율적인 컴퓨터 비전 기반 자동화 관리 시스템의 도입이 필요하다. 본 연구는 공가 통신 설비의 체계적이고 효율적인 관리를 위해 YOLOv8l-seg(You Only Look Once version 8 large-segmentation)으로 객체를 탐지하고, 2DCNN(Convolutional Neural Network)를 통해 두 이미지 간의 객체를 비교하며, 객체의 변화를 감지하는 모델을 구현했다.

2. 관련 연구

2.1 SONet

이 연구는[1] 전력선 점검 작업에서 소 객체 탐지를 위해 YOLOv8 기반인 SONet 알고리즘으로 연구되었다. 이 알고리즘은 SOD(Small Object Detection)과 전력선 점검을 위해 연구되었으며, 탐지 성능으로 mAP(Mean Average Precision)는 78.67%로 평가되었으며, APs(AP for small objects)는 20%로 상대적으로 낮은 성능을 보였다.

2.2 DDCNN

이 연구는[2] 두 시점에서 촬영된 광학 원격감시 이미지의 변화 탐지를 위해 연구되었다. DDCNN은 2DCNN 기반으로 Attention Mechanism과 이미지 차이 증강 기법을 결합하여 두 이미지 간의 차이를 강조하고 변화가 발생한 영역을 픽셀 단위로 탐지한다. 변화 탐지 성능으로 F1 score와 IoU(Intersection over Union) 각각 2.96%, 5.17%로 향상된 성능을 보였다.

3. 본론

본 연구는 YOLOv8l-seg 모델과 2DCNN을 결합하여 객체 변화 탐지 시스템을 설계했다. 데이터는 한국 전봇대 이미지로 Roboflow[4]와 Kaggle[5]에서 수집 후 부족한 이미지는 직접 촬영하여 데이터를 수집했다. Class는 Wire, Crossarm, Pole 세 가지로 정의했다. 전선은 가늘고 복잡한 형태로 이루어져 있어 Polygon 기법으로 Labeling을 진행했다. Labeling 된 이미지는 모델 일반화 성능 향상과 과적합을 방지하기 위해 표 1과 같이 데이터 증강을 하였다. 총 4,475장의 데이터 세트를 구축하였고 Train, Valid, Test는 7:2:1 비율로 사용했다.

증강 기법	적용 범위
Flip	Horizontal
Shear	$\pm 10^\circ$ Horizontal, $\pm 10^\circ$ Vertical
Saturation	Between -25% and +25%

표 1. 이미지 데이터 증강에 사용된 기법 및 적용범위

3.1 YOLOv8l-seg

본 연구는 얇고 복잡한 전선을 탐지하기 위해 YOLOv8l-seg 모델을 사용했다. 전봇대는 전선이 교차하거나 겹친 형태가 많아 다양한 형태와 위치를 파악할 수 있어야 한다. YOLOv8l-seg은 Anchor-free 방식을 사용해 복잡한 전선 환경에서 다양한 형태의 전선을 개별적으로 구별하고 예측할 수 있다. 모델 훈련 하이퍼 파라미터로 batch size:8, epochs:500, image size:640을 사용했다. 모델 평가로는 정량적 평가와 정성적 평가 두 가지 방식으로 진행했으며, 표 2는 정량적 평가 결과로 Crossarm과 Pole Class에서 mAP 0.897, 0.951로 좋은 성능을 보였고, Wire

Class는 0.485로 낮은 수치를 기록하였다. 상대적으로 Crossarm과 Pole의 크기가 크고 명확해 좋은 성능을 보여 주고 Wire와 같이 크기가 작고 가는 객체는 성능이 낮음을 알 수 있다.

Class	precision	recall	mAP50	mAP50-95
All	0.818	0.758	0.777	0.38
Crossarm	0.924	0.863	0.897	0.459
Pole	0.913	0.888	0.951	0.503
Wire	0.617	0.524	0.485	0.179

표 2. YOLOv8l-seg 모델 Instance Segmentation 성능 평가

그림 1은 모델의 Segmentation 예측 결과로 정성적 평가를 진행했다. 건물, 하늘, 나무 등 다양한 배경에서 객체를 탐지하는 것을 확인할 수 있고 특히, 전선이 교차하거나 겹치는 부분에서 전선들을 개별로 잘 구분하는 모습을 알 수 있다.

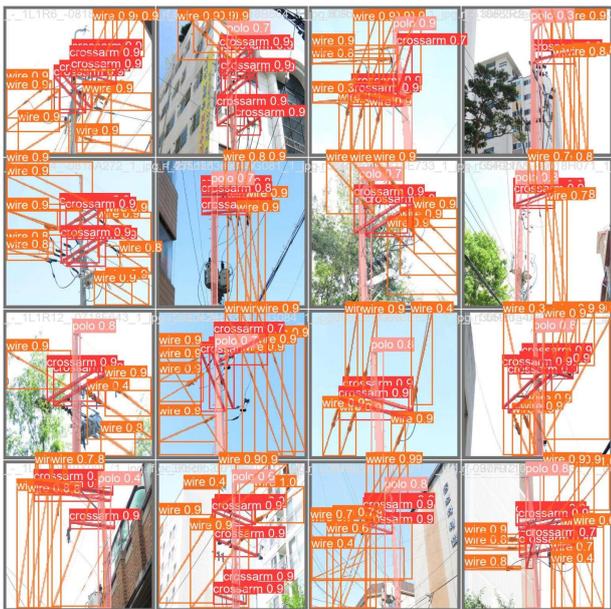


그림 1. YOLOv8l-seg 모델 전선 탐지 및 Segmentation 예측 결과

3.2 2DCNN

두 시점의 이미지를 비교하여 변화를 탐지하는 시스템을 구현했다. 학습된 YOLOv8l-seg의 최적의 모델로 객체를 탐지하고 2DCNN의 유클리드 거리를 사용하여 과거 시점 이미지와 현재 시점 이미지 간의 차이를 계산하고 특징을 비교하여 전선의 변화를 탐지한다. 모델은 두 가지 합성곱층으로 구성되며 첫 번째 합성곱 층에서 입력 이미지의 Edge와 Texture를 추출하고 두 번째 합성곱 층에서 추출된 특징의 Fully Connected Layer를 거쳐 두 이미지 간의 차이를 비교한다. 전선의 변화로 추가는 녹색, 제거는 빨간색으로 표시하며 변화 여부를 시각화한다.

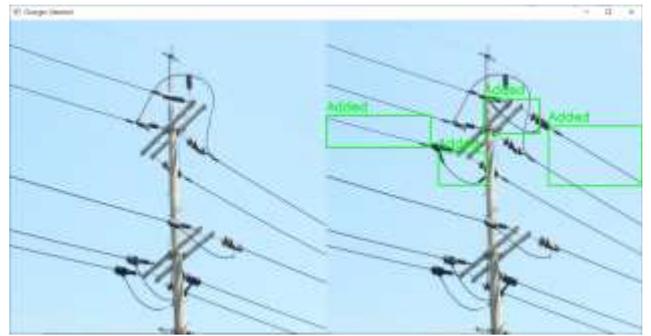


그림 2. 2DCNN을 이용한 객체 변화 탐지 결과

4. 결론

본 연구는 전선의 복잡한 형태와 위치를 탐지하고 전선 변화를 모니터링하기 위해 YOLOv8l-seg 모델과 2DCNN을 결합하여 객체 변화 알고리즘을 설계했다. 연구 결과, Crossarm과 Pole Class에서 준수한 성능을 보였으나, Wire Class에서 상대적으로 낮은 성능을 기록했다. 전선의 가늘고 복잡한 형태가 주요 원인으로 분석되며, 향후 연구는 Wire Class에 대한 탐지 정확도 개선과 다양한 환경, 각도에서 탐지가 가능한 연구가 필요하다.

사사문구

본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음. (2023-0-00028)

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00407739)

참고문헌

[1] The Korea Maritime News (2024). <http://www.haesa-news.com/news/articleView.html?idxno=115165> (accessed SEPTEMBER 01, 2024).

[2] Weicheng Shi, Student Member, IEEE, Xiaoqin Lyu, and Lei Han "SONet: A small object detection network for power line inspection based on YOLOv8", IEEE Transactions on Power Delivery, pp.1-12, AUGUST 2024

[3] Xueli Peng, Herlocker, Ruofei Zhong, ZhenLi, Qingyang Li "Optical Remote Sensing Image Change Detection Based on Attention Mechanism and Image Difference", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. VOL. 59, NO. 9, SEPTEMBER 2021

[4] [Lines] DT Day Tran Dataset V1 Image Dataset(2023). <https://universe.roboflow.com/alysson-machado/-line-s-dt-day-tran-dataset-v1> (accessed March 29 2024).

[5] Telephone Poles Data Set(2019). <https://www.kaggle.com/datasets/chevallier/telephone-poles-dataset> (accessed March 29 2024).

“본 논문은 2024년 한국멀티미디어학회 추계학술대회에 서 발표한 논문을 수정·보완한 것임”

CoD Prompt와 Unsloth를 활용한 미세 조정된 Gemma2 기반의 한국어 문서 요약 효율화

김승주*, 정세훈**, 심춘보*
*국립순천대학교 인공지능공학부
**국립순천대학교 컴퓨터공학과

e-mail : david309@naver.com, cbsim@scnu.ac.kr

Efficiency of Korean Document Summarization based on Gemma2 Fine-Tuned using CoD Prompt and Unsloth

Seung-Ju Kim*, Se-Hoon Jung**, Chun-Bo Sim*

*Dept. of Artificial Intelligence Eng., Suncheon National University

**Dept. of Computer Eng., Suncheon National University

요약

Gemma2 9B-it는 대부분 영어 웹 문서, 수학 및 코드 데이터로 훈련되어 한국어 작업 효율이 떨어진다. 따라서 모델에 한국어 요약 데이터셋을 추가로 학습시키고, Unsloth 라이브러리를 사용해 메모리를 절약하면서 대량의 데이터로 모델을 효율적으로 재학습할 수 있는 CPT 방법으로 미세 조정한다. 또한, 요약문의 정보 밀도를 높이고자 CoD Prompt를 적용하며, 정량적 성능 비교를 위해 ROUGE-N, ROUGE-L, ROUGE-SU, RDASS를 성능 지표로 활용한다. 성능 비교를 위해 미세 조정하지 않은 Gemma2 9B-it와 미세 조정 후 모델의 성능을 비교한다. 실험 결과, 미세 조정된 Gemma2 9B-it 모델이 미세 조정 전 모델보다 한국어 요약 성능이 높았다. 미세 조정된 모델을 활용해 기존 모델보다 효율적인 한국어 요약 서비스를 지원할 수 있고, 보안에 민감한 기관이나 기업에서도 미세 조정된 모델을 자체적으로 호스팅해 안전하게 활용할 수 있을 것으로 기대한다.

1. 서론

일상생활 속에서 방대한 양의 정보를 접하지만 그 속에서 유의미한 정보를 얻기 위해서는 많은 시간과 노력이 필요하다. 방대한 양의 정보를 요약한 문서는 사람들의 정보 획득에 큰 도움을 줄 수 있다.

Gemma2[1]는 미세 조정(Fine-Tuning) 기술을 사용해 특정 작업의 효율을 이전보다 높일 수 있다. Gemma2는 주로 영어 웹 문서, 수학 및 코드 데이터로 구성된 토큰을 통해 훈련되었다.

본 연구에서는 효율적인 한국어 문서 요약을 위해, 모델에 한국어 요약 데이터를 추가로 학습시키고, Unsloth를 사용해 메모리 효율적 미세 조정을 진행했다. 또한, 요약문의 정보 밀도를 높이고자, CoD(Chain of Density) Prompt[2]를 함께 적용한다.

2. 관련 연구

2.1 LoRA(Low-Rank Adaptation)

LLM(Large Language Model)을 특정 작업에 맞게 미세 조정하는 일은 모델의 성능을 효과적으로 개선하는 방법이다. 특히 LoRA[3]와 같은 기술을 활용하면 미세 조정의 효율성이 크게 향상된다. 미리 학습된 가중치를 동결(Freeze)하고 각 계층에 저차원의(Rank-Dificient) 매트릭스를 삽입하여 필요한 파라미터 수를 크게 줄였다. 이를 통해 LLM을 미세 조정할 때 자원과 비용의 부담을 크게 줄여주어 더 많은 연구자와 개발자들이 LLM을 활용할 수

있는 기회를 제공한다.

2.2 CoD Prompting

CoD Prompting은 요약문의 정보 밀도를 점진적으로 증가시키는 방법이다. 초기에는 개체가 적은 요약문을 생성하고, 이후 1~3개의 누락된 개체를 식별하여 추가하는 과정을 반복한다. 이 과정을 5번 반복하면서 개체 대 토큰 비율이 점점 높아지고, 요약문의 밀도가 증가하게 된다. 최종적으로는 사람의 선호도에 부합하는 적절한 밀도의 요약문을 생성하는 것을 목표로 한다.

3. 본론

본 연구에서는 Gemma2 9B(Billion)-it(Instruct)에 한국어 요약 데이터를 추가로 학습시키고, Unsloth를 사용해 컴퓨터 자원을 절약하면서도 대량의 데이터로 모델을 재학습하는 CPT(Continued Pre-Training)를 통해 미세 조정한다. 또한, 미세 조정된 모델이 생성하는 요약문의 정보 밀도를 높이기 위해 CoD Prompt를 적용한다. 성능 비교를 위해 기존 Gemma2 9B-it 모델과 함께, 7B에서 12B 범위의 파라미터 수를 가진 다양한 규모의 모델 간의 정량적 성능을 비교한다.

Gemma2는 대부분 영어로 구성된 데이터로 학습되었기 때문에, 효율적인 한국어 문서 요약을 위해 AI Hub의 한국어 “문서요약 텍스트”, “도서자료 요약” 데이터를 추가로 학습시킨다. 불용어 처리와 Data Filtering을 진행하고, 각 데이터의 Annotation Format을 원문과 요약문으로 정

리한 통합 데이터셋을 구축했다.

LoRA는 기존 모델의 파라미터를 변경하지 않고 특정 모듈만 추가로 학습하여 메모리와 계산 자원을 효율적으로 사용하지만, 학습할 수 있는 파라미터 양이 제한되어 Full Fine-Tuning보다 성능이 낮아질 수 있다. Full Fine-Tuning은 모델의 모든 파라미터를 학습해 대량의 데이터를 완전히 주입하지만, 많은 자원과 시간이 필요하다. Unsloth는 LoRA를 활용해 자원을 적게 사용하면서도 Full Fine-Tuning처럼 대량의 데이터로 모델을 효율적으로 재학습할 수 있는 CPT 방법을 지원한다.

Gemma2 9B-it에 한국어 요약 데이터를 추가로 학습시키고 Unsloth 라이브러리를 활용해 CPT 방법으로 모델을 미세 조정한다. 미세 조정 시 LoRA의 rank는 16, alpha는 64, target_modules는 q_proj, k_proj, v_proj, o_proj로 지정했다. Trainer는 UnslothTrainer를 사용했고, Optimizer는 Adamw_8bit, LR(Learning Rate) 스케줄러는 Cosine 방식을 사용하는 등 Hyper-Parameters를 미세 조정했다.

성능 비교를 위해 미세 조정하지 않은 Gemma2 9B-it를 비교 대상으로 선정했다. 테스트에는 네이버 뉴스 기사 요약 데이터셋[4]을 사용했다.

정량적 성능 비교를 위해 사용한 지표는 ROUGE(Recall Oriented Understudy for Gisting Evaluation)-N, ROUGE-L(Longest Common Subsequence), ROUGE-SU(Skip-Bigram and Unigram), RDASS(Reference and Document Aware Semantic Score)[5]이다. ROUGE-N은 기본적인 어휘와 문맥을 평가한다. 그러나 단독으로 사용하기에는 충분하지 않으므로, ROUGE-L과 ROUGE-SU를 추가로 사용한다. ROUGE-L은 생성한 요약문이 핵심 내용을 잘 유지하고 있는지를 평가하며, ROUGE-SU는 단어의 순서가 아닌 조합으로 성능을 판단하여 한글의 단어 조합과 비연속적 관계를 효과적으로 평가한다. RDASS는 단어의 변형이 빈번한 한글에서 ROUGE 지표만으로 성능을 평가하기 어려운 점을 보완한다. 해당 지표는 요약 모델이 생성한 문장과 실제 정답 문장 간 유사성은 물론, 문서와의 유사성도 함께 파악해 평가한다.

(표 1) 성능 비교

	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-SU	RDASS
FT Gemm2 9B-it	0.3911	0.3173	0.3872	0.2900	0.6603
Gemma2 9B-it	0.3032	0.2408	0.3003	0.2204	0.5474

표 1은 미세 조정된 모델과 미세 조정 전 모델의 성능 지표별 평균 점수를 비교한 결과이다. 제안하는 방법은 모든 성능 지표에서 성능 향상을 나타냈다. 특히, ROUGE-1에서 0.3911, ROUGE-L에서 0.3872의 수치를 보여주며 문장 유사성과 의미적 관련성을 보다 정확히 반영하는 능력이 개선되었음을 확인했다. 또한, 생성된 요약문과 실제 정답의 유사성, 문서와의 유사성을 평가하는 RDASS에서

도 0.6603 수치를 기록하며 전반적인 텍스트 요약 성능이 향상됐다.

4. 결 론

본 연구에서는 Gemma2 9B-it에 한국어 문서 요약 데이터를 추가로 학습시키고, Unsloth를 활용해 CPT 방법으로 모델을 미세 조정했다. 또한, 요약문의 정보 밀도를 높이고자 CoD Prompt를 적용했다. 성능 비교 실험을 통해 미세 조정된 Gemma2 9B-it 모델이 비교 대상 모델들보다 한국어 요약 성능이 높게 나온 것을 확인했다. 따라서 미세 조정된 모델을 통해 기존 모델보다 더 효율적인 한국어 요약 서비스를 지원할 수 있을 것으로 기대한다. 나아가 보안에 민감한 기관이나 기업에서도 제안하는 모델을 자체적으로 호스팅해 안전하게 활용할 수 있을 것으로 사료된다.

향후 양질의 데이터를 추가로 학습시키고, 최적화하여 모델의 성능을 더욱 높이고, 정량 평가뿐만 아니라 사용자 평가와 같은 정성 평가도 추가할 계획이다.

사 사 문 구

본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음. (2023-0-00028)

인 용 문 구

이 연구는 과학기술정보통신부의 재원으로 한국지능정보사회진흥원의 지원을 받아 구축된 "문서요약 텍스트", "도서자료 요약" 데이터셋을 활용하여 수행된 연구입니다. 본 연구에 활용된 데이터는 AI 허브(aihub.or.kr)에서 다운로드 받으실 수 있습니다.

참 고 문 헌

[1] Team, Gemma, et al. "Gemma 2: Improving open language models at a practical size", arXiv preprint arXiv:2408.00118, 2024.
 [2] Adams, Griffin, et al. "From sparse to dense: GPT-4 summarization with chain of density prompting", arXiv preprint arXiv:2309.04269, 2023.
 [3] Hu, Edward J., et al. "Lora: Low-rank adaptation of large language models", arXiv preprint arXiv:2106.09685, 2021.
 [4] Hugging Face(2016). <https://huggingface.co/datasets/daekeun-ml/naver-news-summarization-ko> (accessed Aug., 14, 2024).
 [5] Lee, Dong-yub, et al. "Reference and document aware semantic evaluation methods for Korean language summarization", arXiv preprint arXiv:2005.03510, 2020.

“본 논문은 2024년 한국멀티미디어학회 추계학술대회에서 발표한 논문을 수정·보완한 것임”

동화 대탈출 : 제스처 인식을 활용한 아동 단어 학습 게임

조하진, 김리연, 엄성용
서울여자대학교 디지털미디어학과
email: ha35108@naver.com, fldus4@naver.com, osy@swu.ac.kr

Fairy Tale Escape : A Children's Vocabulary Learning Game Using Gesture Recognition

Ha Jin Cho, Ri Yeon Kim, Seong Yong Ohm
Department of Digital Media, Seoul Women's University

요 약

본 논문은 초등학교 저학년 수준을 타겟으로 한 아동의 단어 학습을 혁신적으로 향상시키기 위해 설계된 PC 기반 학습 게임, <동화 대탈출>을 제안한다. 이 게임은 아동의 읽기 능력을 높이기 위한 목적으로 개발되었으며, 전통적인 학습 방식과 달리 2D 및 3D 환경을 결합하여 학습 경험을 다채롭게 만든다.

게임의 첫 번째 단계에서는 2D 화면에서 각색된 전래동화를 배경으로 숨은 그림 찾기 활동을 통해 단어 학습이 자연스럽게 이루어진다. 이 과정은 아동의 흥미를 유지하고 집중력을 높이며 반복 학습을 돕는다. 두 번째 단계는 3D 환경에서 MediaPipe 기술을 활용한 제스처 인식 기능을 사용하여, 아동이 자신의 손동작으로 캐릭터를 조작하며 단어를 복습할 수 있도록 구성된다. 이러한 감각적 상호작용은 아동이 학습 내용에 더욱 몰입하고, 실제적 경험을 통해 기억을 강화하는 데 도움을 준다.

본 연구에서는 게임의 기능과 설계를 구체적으로 설명하고, 아동의 조기 인터넷 노출에 대한 대안적 학습 도구로서의 가능성을 분석한다. 또한, 본 게임이 아동에게 제공하는 새로운 학습 경험과 그로 인한 교육적 가치를 탐구한다.

1. 서론

최근의 연구에 따르면, 한국 학생들은 국제적으로 읽기 능력에서 낮은 성취도를 보이고 있으며, 특히 장문 독해에서 큰 어려움을 겪고 있다.[1] 2009년과 2018년 사이 OECD의 학업 성취도 비교 결과, 한국은 '축차적 의미 표상'(단어와 문장을 그대로 이해하는 능력)의 정답률에서 15% 이상 감소하여 조사된 5개국 중 가장 큰 하락 폭을 기록하였다. 이러한 통계는 특히 초등학교 저학년 시기의 기초 문해력 발달이 중요함을 시사한다.[2] 따라서, 초등학교 1~3학년과 같은 저학년을 대상으로 한 독해력 교육이 필요함을 알 수 있다.

현재의 교육적 접근은 인공지능(AI) 기술을 활용하여 문해력을 향상시키는 방향으로 발전하고 있다. AI 기반의 문화예술 교육은 이미지 생성 및 기계 학습 등 최신 기술을 활용하여 학습의 효율성을 높이며, 이 과정에서 게이미피케이션(gamification) 기법은 학습자의 동기를 유도하고 몰입감을 높이는 중요한 수단으로 부각되고 있다.

본 연구는 이러한 맥락에서 아동의 언어 학습을 보다 효과적으로 향상시키기 위한 게임형 학습 도구인 <동화 대탈출>을 개발하고, 그 가능성을 탐구한다. <동화 대탈출>은 제스처 인식 기술을 활용하여 아동이 학습에 직접적으로 참여하고, 게임을 통한 즐거운 경험을 통해 단어와 문해력 향상을 도모하는 혁신적인 방안을 제시한다.

2. 관련 연구

게임이 지닌 독특한 특성과 매력은 학습자의 흥미와 참여를 유도하여 학습 동기를 강화하고, 학습 결과에 긍정적인 영향을 미칠 잠재력을 지닌다. 게임 중심의 학습 접근법은 교사의 역할을 재정의하며, 교사를 단순한 게임 활용자나 전달자에서 벗어나 교육 게임을 직접 개발하고 활용하는 창의적인 교육 설계자로 진화시킬 수 있는 가능성을 제시한다.[3] 그 예로, 2012년 노리가 개발한 온라인 수학 게임은 게이미피케이션을 활용해 미국에서 큰 인기를 얻었다.[4]

3. 주요 기능 및 화면 구성

3.1 숨은 그림 찾기를 활용한 단어 학습

본 논문에서 제안하는 시스템의 주요 기능 중의 첫 번째는 숨은그림찾기 게임에서 착안하여 만든 숨은 단어 찾기 기능이다.

그림 1은 화면 속에 숨겨진 단어를 찾는 게임의 기능을 보여준다. 화면 하단에는 학습자가 찾아야 할 단어의 그림이 제시되어 있으며, 사용자는 해당 그림과 일치하는 단어를 화면 속에서 찾아내는 방식으로 게임을 진행한다. 또한 하단에 제시된 단어 그림은 해당 단어의 의미를 제공하여 학습자의 어휘 학습을 돕는다.



(그림 1) 숨은 단어 찾기 예시 화면

3.2 MediaPipe를 사용한 제스처 인식 및 상호작용

본 시스템의 주요 기능 중의 두 번째는 Google의 MediaPipe 프레임워크를 활용한 조작 인터페이스이다. 이 기능은 사용자가 카메라 앞에서 수행하는 손동작을 실시간으로 인식하여, 이를 게임 내 플레이어의 움직임으로 변환하는 과정을 포함한다. 그림 2는 사용자가 카메라에 손동작 모델을 보여주는 모습을 나타내며, 그림 3은 이 손동작에 따라 게임 속 플레이어가 대응하여 움직이는 모습을 보여준다. 그림 2에서 보이는 손동작은 그림 3과 같이 PC 환경에서 시각화되며, 사용자가 손을 움직일 때 플레이어가 이에 즉각적으로 반응하는 장면을 나타낸다.

기능적 원리에 대해 설명한다면 먼저, MediaPipe의 손동작 인식 모듈을 사용하여 사용자의 손 랜드마크(21개의 주요 관절 포인트)를 검출한 뒤 각 관절 사이의 벡터 값을 계산하여 손의 자세를 분석한다. 이를 통해 손의 각도와 위치를 기반으로 한 고차원 특징 벡터를 생성하고, 이를 머신러닝 기반의 제스처 인식 모델에 입력한다. 본 연구에서는 이 과정에서 K-NN 알고리즘을 활용하여 제스처 분류를 수행한다. K-NN 모델은 학습된 데이터셋을 기반으로 각 손동작의 패턴을 학습하며, 거리 기반의 분류를 통해 "Come", "Away", "Spin" 등의 특정 제스처를 실시간으로 인식하도록 설계되었다. 이후 인식된 제스처는 화면 속 가상의 플레이어 캐릭터의 움직임에 매핑된다. 예를 들어, 사용자가 손을 "Come" 제스처로 움직일 경우, 화면 속 캐릭터는 "이리와"라는 동작을 수행하며, "Away" 제스처가 감지될 경우 "저리 가"라는 동작을 수행한다. "Spin" 제스처는 "돌아"라는 동작으로 연결되어 플레이어의 직관적인 동작 인식 및 반응을 가능하게 한다.

이 기능은 단순한 제스처 인식을 넘어, 사용자의 손 움직임 데이터를 실시간으로 분석하고 이를 게임 내 상호작용 요소로 변환함으로써 기존의 마우스나 키보드 입력 방식에서 벗어난 새로운 인터페이스를 제시한다. 이를 통해 사용자 경험을 극대화하고, 게임 플레이의 몰입감을 향상시켜 더욱 효과적인 학습이 이루어질 수 있게 구현하였다.



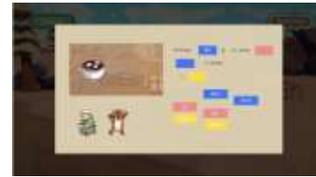
(그림 2) 손 움직임 모델(Come, Away, Spin)



(그림 3) 손 움직임에 따라 플레이어가 움직이는 화면

3.3 이야기 완성을 통한 단어 복습 기능

본 논문에서 제안하는 시스템의 주요 기능 중의 세 번째는 이야기 완성 기능이다. 앞서 숨은 단어 그림 찾기 기능에서 배운 단어들을 활용하여 이야기를 완성한다. 그림 4는 이야기 완성 단계의 기능을 보여준다. 화면의 Story Teller 버튼을 누르면 그림 4와 같이 스토리 구현 화면이 나온다. 빈칸에 알맞은 단어를 넣으면 문장이 완성되면서 이야기가 완성되도록 한다. 이 단어들은 복습을 통해 효과를 더욱더 극대화시켰다.



(그림 4) 이야기 완성

4. 결론 및 향후 연구

본 연구에서는 3D 손 상호작용 기술을 활용한 몰입형 게임 환경 설계를 제안하였다. 게임은 2D 숨은그림찾기와 3D 손 인터랙션 미니게임으로 구성되며, 손 랜드마크 검출과 K-NN 모델을 이용한 제스처 추론 기술을 통해 사용자의 몰입감과 직접적인 상호작용을 이용하여 보다 효과적인 교육용 콘텐츠를 제작하였다. 이를 통해 직관적이고 자연스러운 게임 경험을 제공할 수 있음을 확인했다.

향후, 초등학교 저학년뿐만 아니라 다양한 플레이어 그룹을 대상으로 한 사용자 테스트를 통해 게임 디자인의 실용성을 검증하고, 제스처 인식 정확도를 높이기 위한 추가적인 모델 개선을 계획하고 있다. 또한, 얼굴 촬영 후 캐릭터에 사용자의 얼굴을 인식하여 넣는 기능 등의 다양한 상호작용 요소를 통합하여 보다 풍부한 게임 경험을 제공할 방안을 모색할 예정이다.

ACKNOWLEDGEMENT

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)

참고 문헌

[1] 이우진. "디지털 시대, 읽지 못하는 아이들", 한국교육신문, https://www.hangyo.com/news/article.html?no=98072#google_vignette, 2023.02.03.
 [2] 정중성, 최진오. "초등학교 저학년 아동의 기초 문해력 기준 개발", 『독서연구』 통권 제 70호, 2024. pp173-201
 [3] 김영현. "교육용 게임을 활용한 초등학교 게이미피케이션 사회수업의 교육적 효과 분석", 『한국게임학회 논문지』 Vol.20 No.5, 2020, pp.21-30.
 [4] 염지현 . "게임했는데, 공부가 된다는 '게이미피케이션', 진짜일까?", 동아사이언스, <https://m.dongascience.com/news.php?idx=13362>, 2016.08.10

자율주행 상황 속 환경 객체와 상호작용하는 모빌리티 AR 콘텐츠 개발 방법 연구

박유천, 경병표
국립공주대학교 게임디자인학과
e-mail : 6000@yucheon.io, kyungbp@kongju.ac.kr

A Study on the Development Method of Mobility AR Content Interacting with Environmental Objects in Autonomous Driving Situations

Park Yucheon, Kyung Byungpyo
Dept. of Game Design, Kongju National University

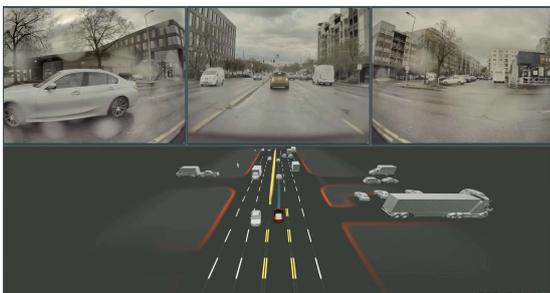
요 약

본 연구는 레벨 4 자율주행 상황에서 외부의 다양한 환경 객체와 상호작용하는 모빌리티 AR 콘텐츠를 제작할 수 있는 개발 환경의 구조와 개발 방법을 제안한다. Unity 엔진을 활용하여 자율주행 시뮬레이터를 제작하고, 구분된 Scene과 API를 기반으로 모빌리티 AR 콘텐츠 개발 환경을 구축하였다. 이를 통해 콘텐츠 제작자는 Unity 엔진으로 AR 콘텐츠를 쉽게 구현할 수 있으며, 콘텐츠 소비자에게 몰입감 있는 AR 경험을 제공할 것으로 기대한다.

1. 서 론

자율주행 기술의 발전으로 많은 완성차 기업에서는 차 내에서 어떻게 콘텐츠를 제공할지에 대한 다양한 콘셉트를 제시하고 있다. 특히 레벨 4 자율주행에서는 시스템이 운영실계영역(ODD) 내의 모든 동적주행작업을 수행하고, 모든 위험 상황에 대비하여 모든 결정을 내리기에 [1], 운전자에게 어떠한 제어도 요구되지 않는다. 이로 인해 자동차 내에서 영화, 게임 등의 콘텐츠 소비도 가능하다.

Tesla는 자율주행을 위해 주행환경을 3D 월드로 변환하는 기술을 개발하였다. 해당 기술과 발전된 AR Heads up display(이하 AR HUD)의 융합으로 자동차 밖의 요소들과 상호작용을 하는 AR 콘텐츠가 전면 유리를 통해 제공될 수 있고, 이러한 콘텐츠를 ‘모빌리티 AR 콘텐츠’라 정의한다.



(그림 1) 주행환경을 인식하여 3D 월드 생성 (Tesla, 2022) [2]

본 논문에서는 미래의 모빌리티 AR 콘텐츠(특히 게임) 제작자가 자율주행 자동차 시스템이 제공하는 정보를 가지고 어떻게 콘텐츠를 개발할 수 있는지 구조를 설계, 개발하고 이를 설명한다.

2. 모빌리티 AR 콘텐츠 제작 환경 구축

2.1 용어 정의

용어	의 미
실제 환경	본 연구의 Unity 시뮬레이션 환경이 아닌 자율주행이 이뤄지는(또는 이뤄질) 실제 환경
환경 객체	주행 환경에서 존재하는 자동차, 건물 등
사용자	자율주행 자동차에서 AR 콘텐츠를 소비하는 사람
사용자 자동차	사용자가 탑승한 자율주행 자동차
콘텐츠 제작자	Unity 엔진을 활용하여 모빌리티 AR 콘텐츠를 제작하는 사람

2.1 시뮬레이션 환경

실제 환경의 자율주행 자동차는 카메라, LiDAR 등의 센서를 활용하여 환경 객체를 검출한다. 본 연구에서는 Unity 엔진을 사용하여 자율주행 환경 시뮬레이터를 제작하였고, 이는 *Environment Scene*에서 처리된다. 자동차, 건물과 같은 환경 객체는 *EnvironmentalObject*(이하 *EnvObj*)로 구현되며, 각 객체의 크기에 맞는 *BoxCollider*를 가진다. 사용자 자동차에는 *SphereCollider*를 추가하여, 충돌한 *EnvObj*를 통해 외부의 환경 객체를 검출한다.

*SphereCollider*의 크기는 실제 환경에서 센서를 통한 환경 객체 검출 가능 범위를 의미한다.

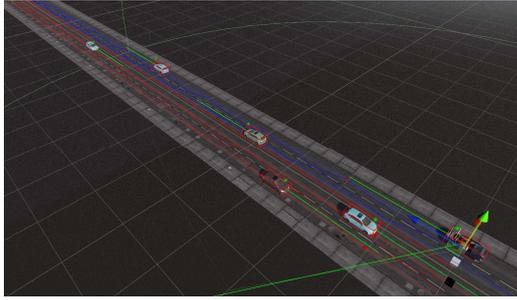
2.2 검출된 세상

검출된 환경 객체는 *Detected World Scene*에서 *DetectedObject*(이하 *DetObj*)로 생성되어 관리되며, 각 *DetObj*는 대응하는 *EnvObj*의 크기와 위치 정보를 포함하고 계속 업데이트한다. 만약 사용자 자동차의 검출 범위를 벗어났을 경우, 해당하는 *DetObj*가 삭제된다.

추가로 사용자 자동차의 위치, 자동차 내 사용자의 머리

및 손 위치 등도 *DetObj*로 제공할 수 있다.

콘텐츠 제작자는 *Detected World Scene* 안에서 자유롭게 Unity를 활용한 개발을 진행할 수 있다.



(그림 2) 검출된 Detected Object (빨간색 박스 테두리 시각화)

2.3 이벤트 API 등을 통한 데이터 획득

사용자 자동차의 객체 검출 시스템(*DetectionManager*)은 관찰자 패턴을 사용하여, *EnvObj* 검출 시작 및 종료 시, *DetObj* 정보를 각 관찰자에게 이벤트로 전달한다. 콘텐츠 제작자는 이를 활용하여 AR 객체를 배치할 수 있다.

```
// 객체 검출 시스템(DetectionManager)에서 새로운 검출된 객체가 생성되었을 때 이벤트를 받는다.
DetectionManager.OnDetectedObjectCreated.AddListener((detectedObject, _) =>
{
    // 몬스터를 생성 위치
    // (자동차 크기의 y만큼 더해서 차 루프 위에 몬스터가 생성될 수 있도록 한다.)
    Vector3 pos = detectedObject.Position; // Detected Object의 pivot은 자동차 최하단 가운데
    pos.y += detectedObject.Scale.y;

    // 몬스터 게임 오브젝트 생성
    GameObject clone = Instantiate(enemyPrefab, pos, Quaternion.identity);

    // 차가 이동해도 몬스터가 차 루프 위에 계속 서 있도록
    // 몬스터의 부모로 Detected Object를 지정한다.
    clone.transform.SetParent(detectedObject.transform);
});
```

(그림 3) 새로 검출된 자동차 위에 AR 몬스터를 스폰하는 스크립트

추가로 사용자 자동차의 속도(*DetectionManager.User CarSpeed*) 등 변수 형태의 정보도 활용할 수 있다.

2.4 실제 환경에서의 구현

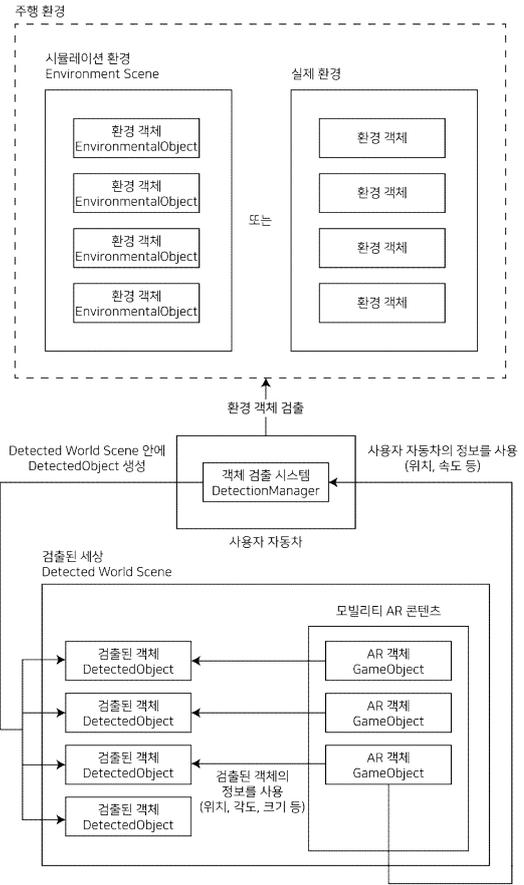
AR HUD로 콘텐츠를 구동할 때, 사용자의 시선과 일치하는 위치에 정확히 AR 그래픽을 출력해야 한다. 이를 위해 Head/Eye Tracking으로 렌더 카메라의 위치를 보정하고, 이후 *Detected World Scene*에서 AR 객체만 렌더링해 AR HUD에 출력한다.

3. 결과



(그림 4) 자동차 위에 몬스터가 생성된 결과

2.3 과정으로 몬스터가 주행 중인 자동차의 *DetObj* 위에 생성되어, 자동차 위에 올라타 있는 것처럼 보인다.



(그림 5) 도식화한 전체 제작 환경 구조

4. 결론

본 연구는 모빌리티 AR 콘텐츠의 제작을 위한 개발 환경과 방법을 제안하였다. 이를 통해 손발이 자유로운 자율주행 환경에서 지속적 운동을 돕는 체감형 게임[3] 등을 AR로 제작하여 이동 중에 플레이할 수 있을 것으로 기대한다.

본 연구는 모빌리티 AR 콘텐츠를 위한 개발 환경의 설계 및 구축에 중점을 두었으며, 향후 연구에서는 이 개발 환경을 활용하여 실제 콘텐츠(게임)를 직접 구현하고, 다양한 응용 가능성에 대해 탐색할 계획이다.

참고 문헌

[1] 이광구, 용부중, 우현구. “레벨 4 자율주행자동차의 기능과 특성 연구”, 『자동차안전학회지』 제12권 제4호, 2020. pp.61-69.
 [2] Tesla. (2022, Oct 1). Tesla AI Day 2022 [Video]. https://www.youtube.com/watch?v=ODSJsviD_SU. (2024년 9월 12일 기준)
 [3] 정석호, 고정운, 허철화, 경병표. “사용자의 헬스케어 위한 AR기반 체감형 게임 콘텐츠 연구”. 『한국게임학회 논문지』 제18권 제4호, 2018. pp.53-63.

ChatGPT를 활용한 인공지능 건강관리 애플리케이션 설계

이호연, 한선아, 정민교
서울여자대학교 소프트웨어융합학과
e-mail : hoit2lee@swu.ac.kr, serah@swu.ac.kr, mchung@swu.ac.kr

Design of AI Health Care Application Using ChatGPT

Hoyeon Lee, Suna Han, Min Gyo Chung
Dept. of Software Convergence, Seoul Women's University

요약

본 논문에서는 ChatGPT를 활용하여 인공지능 건강관리 애플리케이션을 설계한다. 사용자가 자연어를 이용하여 건강 상태를 편리하게 입력할 수 있도록 하고, 사용자에게 미리 제공받은 건강 정보를 바탕으로 개인 맞춤형 피드백을 제공하는 기능을 지원한다. 사용자에게는 피드백과 함께 건강 점수도 제공되고, 이를 바탕으로 일상생활에서 자신의 건강 척도가 어느 정도 수준인지 확인할 수 있다. 제공된 피드백과 건강 점수는 캘린더 형식으로 저장되어 한눈에 건강관리 기록을 확인할 수 있다. 또한 사용자가 자신의 건강 정보를 바탕으로 하여 인공지능 챗봇과의 대화를 진행할 수 있고, 사용자 개인에 맞는 답변을 제공받을 수 있다.

1. 서론

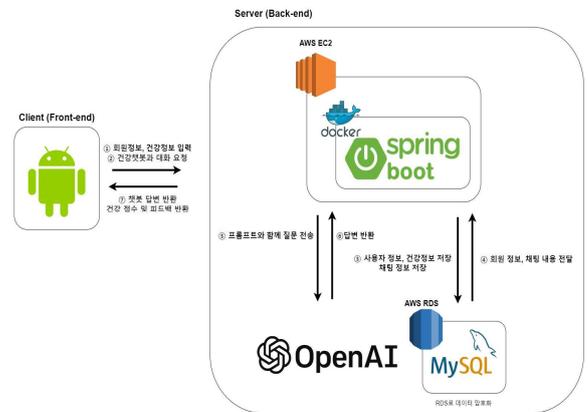
현대 사회에서 인간의 평균 수명이 점점 늘어남에 따라 건강관리는 사람들의 삶에서 매우 중요한 부분으로 대두되고 있다. 특히 코로나바이러스 감염증(COVID-19) 이후로 비대면 서비스의 수요가 크게 증가하였고, 사람들이 원격의료에 대해 가지고 있던 부정적 인식이 완화되면서[1] 디지털 헬스케어 산업의 수요 역시 크게 증가하였다[2]. 또한 20-30대의 젊은 연령층에서 스마트 기기를 사용하여 건강 데이터를 관리하는 사례가 많아지고 있으며, 실제로 건강 서비스 이용 경로에서 스마트폰이 가장 큰 비중을 차지하고 있다[2].

비대면 건강관리 서비스는 단순히 개인 건강관리 측면에서만 그 효과를 발휘하는 것이 아닌, 대면 프로그램의 효과성도 증대시킬 수 있다[3]. 예를 들어, 본 애플리케이션에 기록된 피드백과 건강 점수를 바탕으로 실제 진료에 도움을 줄 수 있다. 운동이나 식단 조절과 같은 물리적인 방법을 이용하여 건강을 관리하는 것도 중요하지만, 현재 자신의 건강 상태를 알고, 기록하여 꾸준히 관리하는 방법도 필요하다. 그러나 현재의 건강관리 애플리케이션은 수치를 입력하고 기록하는 데에만 집중하고 있는 경우가 많다. 매년 사용자가 자신의 데이터를 입력하고 기록하기는 번거롭기 때문에 사용자 입장에서는 일상생활에서 건강을 꾸준히 관리하는 데에 있어 어려움을 느낄 수 있다.

따라서 본 논문에서는 단순히 사용자가 데이터를 입력하고 기록하는 기능에 초점을 맞추기보다는 사용자가 입력한 데이터를 활용하여 ChatGPT가 건강과 관련한 피드백을 사용자에게 전달하고, 건강관리에 대한 궁금증이나 해결책을 챗봇과의 대화를 통해 풀어나갈 수 있는 서비스를 설계한다. 또한 캘린더 형식을 활용하여 자신의 건강관리

기록을 한눈에 살펴볼 수 있도록 한다.

2. 제안 시스템 구성도



(그림 1) 서비스 전체 구성도

그림 1은 본 애플리케이션의 서비스 아키텍처이다. 프론트엔드(Front-End)에서는 사용자가 앱 인터페이스를 통해 건강 관련 데이터를 입력하고, 이를 기반으로 건강 평가 또는 피드백을 요청할 수 있다. 사용자의 입력 데이터는 서버에 전달되며, 백엔드(Back-End)에서는 해당 데이터를 처리하고 관리하는 역할을 담당한다.

백엔드에서는 사용자의 입력 정보와 이전 대화 이력을 데이터베이스에 저장하며, 이를 기반으로 사용자의 건강 상태에 맞는 프롬프트를 생성한다. 이 프롬프트는 OpenAI의 GPT 모델에 전송되어, 사용자의 건강 데이터를 종합

적으로 분석하는 질문 형태로 구성된다.

GPT 모델은 해당 프롬프트에 대한 응답을 생성하며, 필요에 따라 추가적인 정보 요청이나 보완된 프롬프트를 통해 재요청을 수행할 수 있다. 백엔드는 GPT 모델로부터 반환된 응답을 가공하여 사용자가 이해하기 쉽게 프론트엔드로 전달하며, 이와 동시에 사용자에게 맞춤형 피드백을 제공한다.

이를 통해, 시스템은 사용자 맞춤형 건강관리를 지원하며, 사용자는 자신의 건강 상태를 주기적으로 모니터링하고, 즉각적인 피드백을 통해 개선 방안을 제안받을 수 있게 된다.

3. 주요 기능



(그림 2) 애플리케이션 화면 설계

3.1 로그인 및 건강 정보 입력

애플리케이션을 처음 실행할 때, 로그인 화면이 나타난다. 사용자는 로그인 화면에서 아이디와 비밀번호를 입력하여 인증을 진행하며, 이를 통해 사용자 식별이 이루어진다. 인증된 사용자는 이후 재실행 시 로그인을 다시 요구받지 않는다. 아이디와 비밀번호는 서버로 안전하게 전송되며, 회원가입 후 초기 사용자에게는 개인 건강 정보를 추가로 입력하도록 안내한다.

초기 회원 가입 후, 사용자는 자신의 건강 정보를 데이터베이스에 저장하며, 이 정보는 개인의 건강 상태에 적합한 맞춤형 정보를 제공하는 데 활용된다. 개인의 건강 데이터는 저장되어, 사용자의 건강 상태를 기반으로 향후 정보를 활용한 추천이 이루어지도록 한다.

3.2 건강 챗봇 기능

초기 등록 시 데이터베이스에 저장된 사용자의 건강 정보는 종합적으로 분석되어 ChatGPT API를 활용한 건강 챗봇 기능에 통합된다. 이를 통해, 챗봇은 사용자의 건강 상태를 종합적으로 판단할 수 있는 기능을 제공한다. 특히, 프롬프트 엔지니어링 기법을 적용하여, 기존에 수집된 건강 정보를 기반으로 챗봇이 보다 정밀하게 작동할 수 있도록 한다.

구체적으로 사용자의 기저질환이나 건강 상태와 관련된 세부적인 정보를 미리 챗봇에 제공함으로써, 챗봇이 대화 시 보다 정확하고 유의미한 피드백을 제공할 수 있다. 이

러한 접근 방식은 사용자가 챗봇과의 상호작용을 통해 건강 상태를 더욱 효과적으로 모니터링하고, 개인화된 건강관리 조언을 받을 수 있도록 지원한다. 이로 인해, 챗봇은 사용자의 건강 정보에 기반한 맞춤형 대화가 가능해져 더욱 심층적이고, 신뢰성 높은 건강관리 서비스를 제공할 수 있다.

3.3 건강 점수 및 피드백 로직

사용자는 일일 활동량, 수면 시간, 스트레스 수준 등의 데이터를 입력하며, 각 항목에 가중치를 부여해 건강 상태를 평가한다. 예를 들어, 수면 시간은 점수의 30%, 활동량은 20%를 차지하도록 설정된다. 이 평가를 통해 0점에서 100점 사이의 건강 점수가 산출되며, 사용자의 건강 상태를 쉽게 이해할 수 있는 지표로 제공된다.

예를 들어, 하루 평균 수면 시간이 5시간, 일일 걸음 수가 3,000보인 경우, 건강 점수는 낮게 책정되며 추가적인 조치가 필요하다는 경고를 받는다. 애플리케이션은 이를 바탕으로 맞춤형 피드백을 제공하며, "수면 시간이 부족합니다. 일찍 잠자리에 드세요."와 같은 구체적인 권고 메시지를 통해 사용자의 건강 개선을 지원한다.

4. 결 론

본 논문은 ChatGPT를 활용하여 인공지능 건강관리 애플리케이션을 설계하였다. 사용자의 일상 건강 데이터를 기반으로 건강 점수를 계산하고 피드백을 주는 기능과 사용자가 특정 증상이나 궁금증을 입력하면 챗봇이 이를 분석하여 대답하는 기능으로 구성되었다. 위 기능으로 사용자는 자신의 기록을 꾸준하게 관리하는 데 용이하고, 이는 의료 분야 같은 실생활에서 유용하게 활용되는 효과를 기대할 수 있다.

현재 Android 애플리케이션 개발을 진행 중이다. 애플리케이션을 배포한 후에는 애플리케이션 화면 편의성, 기능의 유용성, 챗봇 응답 내용의 신뢰성, 타 애플리케이션과의 차별성, 총 4가지 항목을 바탕으로 사용자 성능 평가를 진행할 예정이고, 이를 바탕으로 애플리케이션의 성능을 개선시킬 예정이다. 많은 사용자의 편리한 건강관리에 도움이 되기를 기대한다.

Acknowledgement

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음 (2024)

참 고 문 헌

[1] 김지예, 이지영, 양선영, 김혜경. “원저 디지털 기반의 생애주기별 건강관리서비스의 핵심 영역”, 『보건교육건강증진학회지』 제40권 제4호, 2023, pp.33-43.
 [2] 오유미. “종설 비대면 건강관리서비스 현황과 발전방향”, 『보건교육건강증진학회지』, 제40권 제4호, 2023, pp.21-32.
 [3] Choi Eunjin. “Health Impact Assessment on Digital-Based, Non-Contact Healthcare Service”, 한국보건사회연구원, 2022, pp.289.

생성형 AI를 활용한 게임 시나리오 생성에서 사전 학습과 계층적 프롬프팅이 사용자 경험에 미치는 영향

허주희, 이진주, 이소정, 정민교
서울여자대학교 소프트웨어융합학과
e-mail : mooo0502@swu.ac.kr, roskf1000408@naver.com, sj011105@swu.ac.kr, mchung@swu.ac.kr

The Effect of Pre-learning and Hierarchical Prompting on User Experience in Generative AI-based Game Scenario Generation

Juhui Heo, Jinju Lee, Sojeong Lee, Min Gyo Chung
Department of Software Convergence, Seoul Women's University

요약

본 연구는 생성형 AI를 활용한 시나리오 생성에서 사전 학습과 계층적 프롬프팅이 사용자 경험에 미치는 영향을 분석한다. 연구는 사전 학습과 계층적 프롬프팅을 함께 사용할 경우 사용자 경험이 더 긍정적일 것이라는 가설 1과, 두 요소가 모두 사용되지 않을 경우 부정적 평가를 받을 것이라는 가설 2를 수립하였다. 15명의 피험자를 대상으로 4가지 실험 조건을 적용하여 사용자 경험을 평가한 결과, 사전 학습과 계층적 프롬프팅을 함께 사용할 때 일부 항목에서 유의미한 긍정적 평가가 나타났다. 특히 계층적 프롬프팅만으로도 사용자 경험을 향상시킬 수 있음을 확인하였으며, 이는 향후 생성형 AI 기반 시나리오 개발에서 중요한 시사점을 제공한다.

1. 서론

게임에서 시나리오는 사용자 몰입도와 체험 지속성을 결정하는 중요한 요소 중 하나이다. 시나리오를 작성할 때는 기존 사용자의 요구와 독창성을 모두 고려해야 한다. 이를 위해 생성형 인공지능(Artificial intelligence, AI)은 기존 데이터를 분석하고 사용자 행동을 학습하여 빠르게 시나리오를 생성하는 도구로 주목받고 있다. 이러한 AI 기반 시나리오 생성은 몰입도를 높이고, 사용자의 선택에 따라 경험을 맞춤화하여 리플레이 가능성과 몰입성을 증대시킬 수 있다. 이러한 AI 활용 사례는 교육, 게임, 의료 등 다양한 분야에서 맞춤형 경험을 제공하는 데 효과적이라는 점이 강조되고 있다 [1, 2].

여기에 더해 생성형 AI는 추가적인 기법을 활용하여 더욱 효율적인 결과를 도출할 수 있다. '사전 학습'을 통해 기존 데이터를 분석하고 학습하면 상용 서비스의 장점과 사용자 특성을 반영할 수 있어 유용하다. '계층적 프롬프팅'은 기승전결 구조에 따라 기본 줄거리에서 세부적인 시나리오를 작성하게 해주며, 이를 통해 더욱 자연스럽게 일관된 스토리라인을 생성할 수 있다.

따라서, 본 연구는 생성형 AI에 '사전 학습'과 '계층적 프롬프팅'을 도입하여 시나리오를 생성하고, 생성된 시나리오가 사용자의 주관적 경험에 미치는 영향을 분석하는 것을 목표로 한다. 본 연구는 두 가설을 수립했다. 사전 학습과 계층적 프롬프팅을 모두 사용할 경우 긍정적 사용자 경험이 가능하다 (가설 1). 사전 학습과 계층적 프롬프팅이 모두 사

용되지 않는 경우 사용자 평가가 가장 부정적이다 (가설 2).

2. 생성형 AI를 활용한 시나리오 생성 방법

시나리오 제작을 위해 다양한 생성형 AI 모델을 비교한 결과, 최종적으로 Chat GPT-4o를 선택하였다. 이 모델은 대화형 인터페이스와 빠른 속도, 피드백 및 수정이 용이하다는 점에서 장점이 있었다.

사전 학습이 있을 경우, 나무위키에서 크롤링한 '스타듀 밸리' 시나리오 데이터를 전처리하여 학습 데이터로 사용하였다. 이를 바탕으로 시나리오 생성을 요청했으며, 사전 학습이 없는 경우에는 단순히 시나리오 생성 프롬프트만 입력하였다. 계층적 프롬프팅은 기본적인 시나리오 생성 후, 추가적으로 세부적인 서술을 요구하는 방식으로 진행되었다.

3. 사용자 평가

제작된 시나리오를 평가하기 위해 피험자들은 네 조건에서 평가를 진행했다. 조건 1은 사전 학습과 계층적 프롬프팅을 모두 사용한 경우, 조건 2는 사전 학습만 사용, 조건 3은 계층적 프롬프팅만 사용, 조건 4는 두 가지 모두 사용하지 않은 경우로 구성되었다.

3.1 피험자

피험자는 시나리오 기반 게임을 최소 한 번 이상 플레이한 평균 연령 25.87세(표준편차: 2.16)의 성인 남녀를 모집했으며, 자발적 동의를 통해 선정되었다. 결과적으로 총 15명

(남성 8명, 여성 7명)의 피험자가 모집되었다.

3.2 사용자 경험 평가

시나리오를 평가하는데 User Experience Questionnaire (UEQ) 설문지를 사용하였다. UEQ는 다음의 항목을 포함한다. 어려움↔쉬움(항목 1), 복잡함↔간단함(항목 2), 비효율적↔효율적(항목 3), 명확하지 않음↔명확함(항목 4), 지루함↔흥미로움(항목 5), 재미없음↔재미있음(항목 6), 불편함↔편안함(항목 7), 동적↔모던/동적(항목 8).

3.3 실험 과제

시나리오는 평가의 효율성을 위해 줄글 형태에 한 가지 스토리 라인으로 제시되었다. 피험자는 종이에 프린트된 시나리오를 읽고 평가를 진행했다.

3.4 실험 절차

모든 피험자는 네 가지 실험 조건을 무작위 순서로 체험하고, UEQ 설문지를 통해 각 조건에 대한 만족도를 평가했다. 실험 전, 절차와 조건에 대한 간략한 설명을 들었으며, 조건 차이에 대한 추가 정보는 제공되지 않았다. 시나리오 체험 후 즉시 설문 평가가 이루어졌고, 피로도를 고려해 각 설문 후 적절한 휴식 시간이 주어졌다. 또한, 학습 효과를 방지하기 위해 순서는 상쇄균형화(Counterbalancing) 방식으로 조정되었다.

4. 연구 분석 결과

네 가지 조건에 따른 결과는 다음과 같다 (그림 1 참조). 조건 1에서는 모든 항목에서 긍정적인 평가를 받았고, 항목 4에서 +2.2, 항목 3에서 +1.9의 높은 점수가 기록되었다. 조건 2에서는 조건 1보다 전반적으로 낮은 평가를 받았으며, 항목 2에서 +1.9, 항목 1에서 +1.7의 점수를 기록한 반면, 항목 7에서 -1.0, 항목 5에서 -0.6으로 부정적인 평가가 나왔다. 조건 3에서는 항목 1에서 +2.1, 항목 2에서 +2.5로 가장 높은 점수를 기록했으나, 항목 5에서 -0.7로 부정적인 평가가 나왔다. 조건 4에서는 전반적으로 부정적인 평가를 받았으며, 항목 7에서 -2.1로 가장 낮은 점수가 기록되었다.

5. 결론

5.1 가설 검증 및 논의

본 연구는 사진 학습과 계층적 프롬프팅이 사용자 경험에 미치는 영향을 분석하였다.

가설 1에 따르면, 사진 학습과 계층적 프롬프팅을 동시에 사용할 경우 항목 3, 항목 5, 항목 6에서 긍정적인 평가를 받았으며, 통계적으로 유의미한 차이가 나타났다($p\text{-value} < 0.05$). 하지만 다른 항목에서는 유의미한 차이가 나타나지 않아, 가설 1은 부분적으로만 검증되었다. 가설 2에서는 사진 학습과 계층적 프롬프팅이 모두 사용되지 않을 경우 항목 4에서 부정적인 평가가 나왔으나($p\text{-value} < 0.05$), 다른 항목들에서는 유의미한 차이가 나타나지 않았다. 따라서 가설 2도 부분적으로만 검증되었다. 또한, 조건 3에서 항목 1과 2에서 각각 +2.1과 +2.5의 높은 점수를 기록한 것은 계층적 프롬프팅만으로도 사용자가 시나리오를 쉽게 이해하고 간단하게 느낄 수 있음을 시사한다. 이는 사진 학습 없이도 프롬프팅 방식이 사용자 경험에 중요한 영향을 미칠 수 있음을 보여준다.

5.2 연구 의의 및 향후 적용

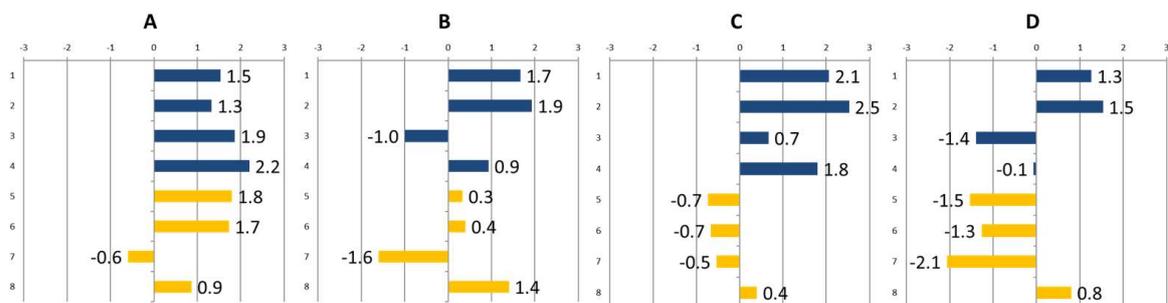
본 연구는 사진 학습과 계층적 프롬프팅이 사용자 경험에 긍정적인 영향을 미칠 수 있는 중요한 요소임을 확인하였으며, 특히 시나리오의 구조적 명확성과 사용자 친화성을 개선하는 데 중요한 역할을 한다는 점을 강조한다. 이는 향후 생성형 AI를 활용한 시나리오 제작에서 사용자 경험을 더욱 향상시키기 위한 중요한 기초 자료가 될 것이다.

5.3 한계 및 향후 연구 방향

본 연구는 게임 시나리오 생성에서 생성형 AI의 사진 학습과 계층적 프롬프팅이 사용자 경험에 미치는 영향을 실험적으로 분석하였으나, 몇 가지 한계점이 존재한다. 첫째, 실험에 참여한 피험자의 표본 수가 15명으로 제한적이었으며, 이는 결과의 일반화 가능성에 한계를 줄 수 있다. 보다 다양한 연령층과 배경을 가진 더 많은 피험자를 대상으로 추가적인 연구가 필요하다. 둘째, 일부 항목에서 통계적으로 유의미한 차이가 나타나지 않은 이유에 대한 추가적인 분석도 필요하다. 특히, 계층적 프롬프팅만 사용한 조건에서 몇몇 항목에서 높은 평가를 받았으나, 다른 항목에서는 기대보다 낮은 평가가 나왔다. 이는 프롬프팅 방식의 구체적인 설계와 적용 방식에 따라 사용자 경험이 달라질 수 있음을 시사하며, 이를 최적화하기 위한 구체적인 연구가 필요하다.

감사의 글

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)



(그림 1) A, B, C, D는 각각 조건 1, 2, 3, 4의 UEQ 항목별 평균값

참고 문헌

- [1] Xu Guo, Yiqiang Chen. "Generative AI for Synthetic Data Generation: Methods, Challenges and the Future." Nanyang Technological University and Institute of Computing Technology, Chinese Academy of Sciences.
- [2] Meng Li, Yaning Li, Chenyang He, Houxiang Wang, Jiaqian Zhong, Shengxin Jiang, Mingtao He, Zhining Qiao, Jiawang Chen, Yuan Yin, Ray Lc, Ji Han, Zengyao Yang, Mohammad Shidujaman. "Generative AI for Sustainable Design: A Case Study in Design Education Practices." Human-Computer Interaction (HCI 2024), Conference paper, First Online: 01 June 2024, pp. 59 - 78.

생성형 AI 기반의 시험문제와 핵심 요약 제작 서비스

이채영, 김민영, 정민교
서울여자대학교 소프트웨어융합학과
e-mail : 0929alice@naver.com, rucy0716@naver.com, mchung@swu.ac.kr

Generative AI-Based Test Question and Key Summary Creation Service

Chaeyoung Lee, Minyoung Kim, Min Gyo Chung
Dept. of Software Convergence, Seoul Women's University

요약

본 논문은 AI(Artificial Intelligence) 기반의 학습 도구인 'MAP(My AI Professor)'을 제안한다. 이 시스템은 AI를 통해 강의 자료 내용을 바탕으로 예상 문제 및 핵심 요약을 생성하는 기능을 제공한다. 이를 위해 OpenAI의 GPT(Generative Pretrained Transformer)-3.5를 사용하여 강의 자료의 내용을 바탕으로 문제를 생성하고, 핵심 요약서를 제공하는 기능을 개발한다. 시험문제의 정확성을 판단하기 위해 중복성, 조건 충족 여부에 대해 실험하였고, 핵심 요약서의 사용성을 향상하기 위해, 대학생들을 대상으로 사용성 평가를 진행하였다. 이를 통해 MAP의 효과적인 학습 도구로서의 가능성을 발견할 수 있었다.

1. 서론

최근 생성형 AI 분야에 대한 관심이 높아지고 있다. ChatGPT와 같은 생성형 AI 서비스들은 사용자의 질문에 따라 공학적 문제뿐만 아니라 미학, 문학, 철학 등 다양한 분야에서 창의적이고 전문적인 답변을 생성한다. 이러한 생성형 AI 시대에 접어들며 산업, 공공, 노동에도 다양한 변화가 나타나고 있다[1].

교육 분야에서도 AI 시스템을 적극적으로 도입하는 추세를 확인할 수 있다. 첨단 테크놀로지의 발달은 교수학습 활동의 변화를 가져온다. 테크놀로지는 새로운 지식을 구성하고, 문제를 해결하고, 다른 학습자와 상호작용하는 방식에 영향을 미치기 때문이다[2]. 교육의 디지털화에 따라 학생들은 효율적으로 학습 자료를 활용하고 있다. 자동화된 문제 출제는 학습자들에게 다채로운 학습 방식을 경험하고, 학습 효율을 높일 수 있도록 돕는다. 그러나 최근 국내 AI 활용 교육 연구 동향에 따르면, '학생 교육' 유형의 AI 활용시스템은 23.1%에 불과하다[3]. 이에 따라 본 연구에서는 대학생을 위한, 강의 자료를 바탕으로 예상 문제 및 핵심 요약서를 제공하는 학습 도구인 MAP(My AI Professor)을 제안한다. 본 연구의 최종 목표는 MAP 서비스가 학습자들에게 어떤 가치를 제공하는지를 밝히고, 생성형 AI의 가능성을 탐구하는 것이다.

2. 관련 연구

2.1 생성형 AI

생성형 AI는 대규모 언어 모델(Large Language Model)을 적용하여 마치 인간처럼 질문에 답변하는 인공지능의 한 형태이다[4]. 대표적인 예로 ChatGPT가 있으며, 이 기술은 논문 작성, 프로그래밍, 문서 요약 등 다양

한 영역에서 활용되고 있다.

2.2 프롬프트 엔지니어링

프롬프트 엔지니어링은 AI 모델에게 특정한 지시어(프롬프트)를 입력하여 원하는 출력을 얻어내는 기술이다. 프롬프트는 결과 도출을 위한 명확한 '지시', AI가 맥락 및 상황에 따라 반응하도록 하는 '조건' 또는 '역할', 결과 생성 시 필요한 데이터 '입력 값'과 '결과 출력 형식' 등으로 구성된다[5]. 프롬프트의 작성 방식에 따라서 답변의 정확도가 바뀔 수 있으며, 정확한 답변을 얻기 위해서는 프롬프트를 정교하게 설계해야 한다[6].

3. 시스템 설계 및 구현

3.1 서비스 실행 프로세스

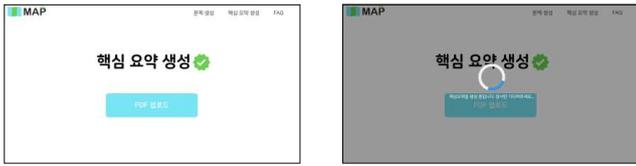
MAP은 강의 자료를 업로드하여 예상 문제와 핵심 요약을 자동으로 생성하는 시스템이다. OpenAI의 GPT-3.5를 활용하여 문제 생성 및 요약서를 제공한다. 사용자는 문제 유형과 문제 수를 선택하여 객관식, 단답형, 서술형 문제를 생성할 수 있으며, 생성된 문제와 요약서는 PDF 형식으로 다운로드 가능하다.



(그림 1) 실제 웹사이트 화면 : 홈



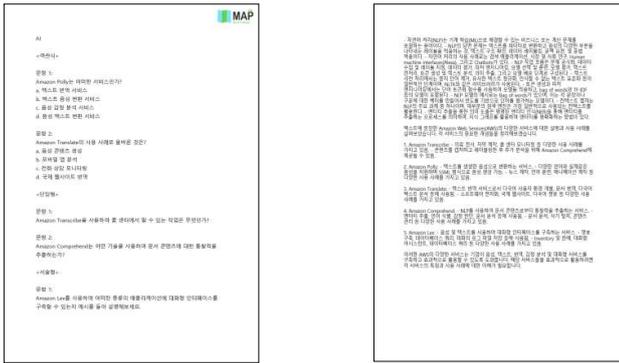
(그림 2) 실제 웹사이트 화면 : 문제 생성 / 문제 생성 중



(그림 3) 실제 웹사이트 화면 : 핵심 요약 생성 / 핵심 요약 생성 중



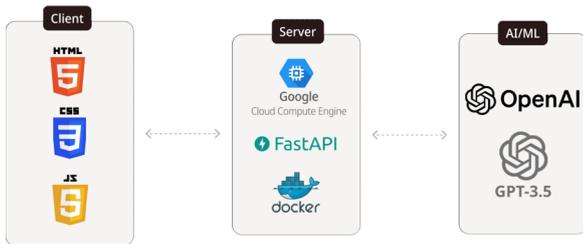
(그림 4) 실제 웹사이트 화면 : 생성 완료



(그림 5) 실제 생성된 예시 : 문제 생성 / 핵심 요약 생성

3.2 MAP의 시스템 구성과 모듈

MAP은 그림 6처럼 PDF 자료를 텍스트로 변환해서 GPT-3.5로 넘긴 후, 그 결과를 웹 서비스로 확인할 수 있게 구현한다. 사용자가 학습한 내용을 자체적으로 점검해 볼 수 있도록 자동으로 생성된 핵심 요약과 예상 문제를 다운로드할 수 있는 시스템으로 구성되어 있다. 웹은 html, css, js로 구현하였고, server는 Google의 Cloud Compute Engine과 FastAPI, docker를 활용해 구현하였으며, AI/ML 부분은 OpenAI의 GPT-3.5를 활용하였다.



(그림 6) 시스템 개요

4. 실험

4.1 실험 개요

본 실험은 앞서 제안한 생성형 AI 기반의 시험문제와 핵심 요약 제작 서비스 MAP의 정확성을 판단하고 사용성을 향상하는 것을 목적으로 한다. 생성된 시험문제지의 정확성을 판단하기 위해 문제들의 중복성과 조건 충족 여부를 측정한다. 중복성은 생성된 문항들 사이에 중복된 내용이 존재하는지를 확인한다. 조건 충족 여부는 사용자가 선택한 문제 유형별로, 입력한 문항 수의 문제를 제대로 생성하는지 확인한다. 생성된 핵심 요약의 정확성, 효율성, 가독성 그리고 만족도에 대한 질적 데이터를 측정하기 위

해 사용성 평가를 수행한다.

4.2 문제 생성 정확성 측정 실험

총 50개의 시험지를 생성하고, 중복성 및 조건 충족 여부를 평가하였다. 중복률은 26%, 조건충족률은 76%로 나타났다.

(표 1) 문제 생성 정확성 측정 실험 결과

측정 항목	실험 횟수
중복성	13 / 50
조건 충족	38 / 50

4.3 핵심 요약서 사용성 평가

사용성 평가는 대학생 22명을 대상으로 진행되었으며, 정확성 4.59점, 가독성 3.27점, 효율성 4.04점, 만족도 4점으로 평가되었다.

(표 2) 사용성 평가 결과

설문 문항	평균 점수
핵심 요약서의 내용이 강의자료와 비교 해봤을 때, 정확하다.	4.59
핵심 요약서의 가독성이 좋다.	3.27
핵심 요약서가 학습 효율을 높이는 데 도움이 된다.	4.04
핵심 요약서를 생성하는 인공지능 서비스를 사용할 의향이 있다.	4

5. 결론

본 연구에서 제안한 MAP 서비스는 대학생들의 학습 효율성을 높이는 데 긍정적인 효과가 있었다. 다만, 문제 생성의 중복률을 줄이고 요약서의 가독성을 개선할 필요가 있다.

Acknowledgement

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구 결과로 수행되었음 (2024)

참고문헌

[1]김태원. (2023). 챗GPT가 촉발한 생성형 AI 시대, 미래 모습과 대응 방안. FUTURE HORIZON,(55), 2-9.
 [2] 조영환, 이지은, 임규연, 정한호, 한인숙. (2023). 생성형 AI와 함께하는 미래교육: 기계에서 협력 파트너로. 교육공학연구, 39(4), 1449-1478
 [3] 허미선, 배윤주, 석희진, 이정민. (2021). 국내 AI활용 교육 연구동향. 정보교육학회논문지, 25(6), 973-985.
 [4] 한정훈. (2023). 생성형 AI 시대의 개막. 한국방송통신전파진흥원.
 [5] 변공규, 최권택, 유선진. (2023). 생성형 인공지능을 이용한 질문 생성 기반 AI 어시스턴트 북. 한국정보기술학회논문지, 21(12), 217-226, 10.14801/jkiit.2023.21.12.217
 [6] 박상언, 강주영. (2023). ChatGPT 및 거대언어모델의 추론 능력 향상을 위한 프롬프트 엔지니어링 방법론 및 연구 현황 분석. 지능정보연구, 29(4), 287-308.

정보화 솔루션을 사용하는 생산현장 내 근무특성을 고려한 음성과 모션 Data 기술

장정우*, 하호수**, 임혜나**, 조훈희*, 김용현***, 김예진***

*전북대학교 컴퓨터공학과

**전북대학교 IT정보공학과

***(주)블루비즈

email : jeongwoo1998@daum.net*, hasoma@naver.com**,

h_naa@jbnu.ac.kr**, 129johh@naver.com*,

yhkimway@naver.com***, yejinoo@blue-biz.co.kr***

Voice and Motion Data technology considering the characteristics of work in the production site using informatization solutions

Jeong-Woo Chang*, Ho-su Ha**, Hae-Na Lim**, Hoon-Hee Cho*,

Yong-Hyun Kim***, Ye-Jin Kim***

*Dept of Computer Engineering, Jeonbuk National University

**IT Information and Engineering, Jeonbuk National University

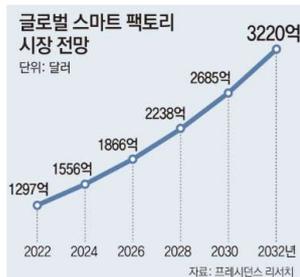
***BlueBiz Inc.

요 약

스마트 공장은 4차 산업혁명의 핵심 기술을 공장과 접목하여 생산 시스템을 자동화 및 지능화하는 혁신적인 생산 방식으로, 생산성 향상과 안정성 강화 등 다양한 장점을 가지고 있다. 이에 따라 최근 스마트 공장은 점점 더 늘어나고 있는 추세이다. 하지만, 현재 스마트 HACCP(Hazard Analysis and Critical Control Points) 솔루션을 사용하는 과정에서 근로자들의 손이 오염되어 있어서 직접 손으로 데이터를 입력하기 어려운 부분이 있다. 본 논문에서는 Speech Recognition API를 이용한 STT(Speech To Text) 기술을 통해 음성 인식 기술을, MediaPipe 라이브러리와 학습시킨 모델을 통해 모션 인식 기술을 개발하고, 이를 HACCP 솔루션 서버와 연동하여 물리적인 접촉 없이 음성과 모션 인식을 통해 데이터를 처리할 수 있는 기술을 제안한다.

1. 서 론

최근 들어 전반적인 생산 시스템을 자동화하고, 지능화한 스마트 공장 구축이 증가하고 있는 추세이다.



(그림 1) 스마트 공장 성장 추이[1]

이러한 스마트 공장에서는 스마트 HACCP 솔루션을 사용하여 공장의 전반적인 데이터를 관리하고 있다. 스마트 HACCP 솔루션을 사용하기 위해서는 손으로 직접 데이터를 입력해야 하는데, 생산 현장의 근로자들은 손이 오염되어 있는 경우가 많기 때문에 데이터를 입력할 때 불편을 겪는 경우가 많다. 따라서 직접 손으로 데이터를 입력하는 방식이 아닌, 음성과 모션 인식 기술을 통해 사용자

의 음성과 모션으로 데이터를 입력하거나 처리할 수 있는 방식을 개발하여 실제 생산 현장에서 근로자들이 데이터를 편리한 방식으로 입력하거나 처리할 수 있는 방식을 보급하고자 한다.

2. 관련 연구

2.1 OpenCV

OpenCV은 Open Source Computer Vision의 약자로, 실시간 이미지 및 동영상 처리를 목적으로 사용되는 컴퓨터 비전 라이브러리이다. 이후 카메라를 통해 모션 인식에 필요한 동작을 인식하고, 추적하는 데 사용한다.

2.2 MediaPipe 라이브러리

MediaPipe 라이브러리는 Google에서 개발한 오픈 소스 라이브러리로, 실시간으로 영상과 이미지 데이터를 처리 및 분석하는 데 사용된다. 이는 임의의 감각 데이터에 대한 추론을 수행하기 위한 파이프라인 구축 프레임워크이다[2]. 본 논문에서는 MediaPipe Hands를 사용하여 각 손가락 마디에 있는 21개의 랜드마크를 통해 손 모양을 추적한다.



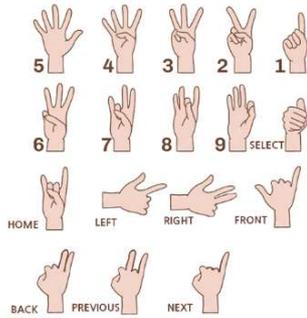
(그림 2) MediaPipe 3D 랜드마크

2.3 데이터 셋 정의

스마트 HACCP 솔루션을 사용하는 데 필요한 음성 및 모션 인식 기술 데이터 셋을 정의한다.

음성 인식의 경우에는 STT(Speech To Text) 기술을 사용하기 때문에 명령에 필요한 단어 및 문장을 정의한다.

모션 인식의 경우에는 손 모양을 이용하기 때문에 명령에 필요한 손 모양 데이터 셋을 정의한다.



(그림 3) 모션 인식 데이터 셋 정의

2.4 Speech Recognition API

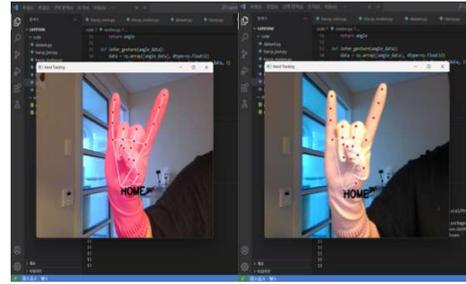
음성 인식 기술은 자동적 수단에 의하여 음성으로부터 언어적 의미 내용을 식별하는 기술로, 마이크와 같은 소리 입력 센서를 통해 얻은 음향학적 신호의 특징을 추출하여 음성 모델 데이터베이스와 비교하는 과정을 통해 음성 인식을 하게 된다[3].

본 논문에서는 Speech Recognition API를 이용한 STT(Speech To Text) 기술을 통해 마이크로 입력된 음성 명령어를 문장으로 변환한다. 이후 문장으로 변환된 음성 명령어와 사전에 정의한 데이터 셋을 비교하여 일치하면 해당하는 동작을 수행할 수 있도록 한다.

2.5 KNN 모델 학습

모션 인식 모델을 학습시키기 위해서 MediaPipe를 사용하여 이전에 정의한 데이터 셋의 손 모양마다 손가락마다 사이 각도를 계산하고, 이를 .csv 파일로 저장하여 데이터 셋을 구축한다.

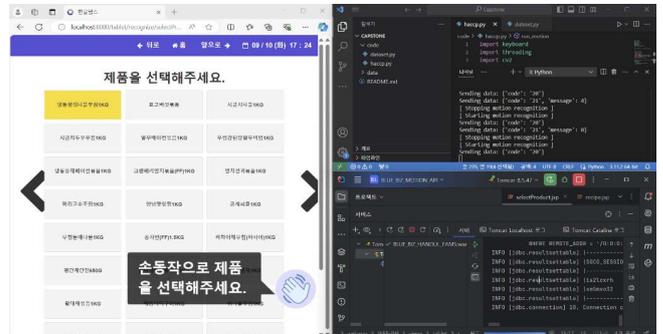
이후 구축한 데이터 셋을 통해 KNN 모델을 학습시킨다. KNN 모델은 K-Nearest Neighbor의 준말로, 새로운 데이터가 들어왔을 때 K개의 가장 가까운 데이터들과 비교하여 예측하고 분류하는 알고리즘이다. 사전에 구축한 손 모양 레이블과 매핑을 통해 학습한 모델이 예측을 수행할 수 있도록 한다.



(그림 4) KNN 모델을 통한 모션 인식

2.6 음성 및 모션 인식 기술을 연동한 HACCP 솔루션 개발

기존의 HACCP 솔루션 서버에 음성 및 모션 인식을 통해 처리된 데이터를 WebSocket 방식으로 전달하여 HACCP 솔루션을 조작할 수 있도록 한다. 이를 통해 직접 손으로 데이터를 입력할 필요 없이 음성과 모션으로 HACCP 솔루션을 조작할 수 있게 된다.



(그림 5) 실제 사용 화면

3. 결론

실제 생산 현장에서는 근로자들이 스마트 HACCP 솔루션 사용 시 손이 오염되어 있어서 직접 손으로 데이터를 입력하거나 처리하기 어려운 상황에 처해있다. 이를 극복하기 위해서 본 논문에서는 음성 및 모션 인식 기술을 개발하여, 기존의 스마트 HACCP 솔루션 서버와 연동하였다. 이를 통해 사용자는 물리적인 접촉 없이, 음성과 모션을 통해 데이터를 입력 및 처리하여 스마트 HACCP 솔루션을 이용할 수 있게 된다.

사 사

“본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 지원을 받아 수행되었음” (2022-0-01067)

참고 문헌

- [1] 박현익, "새로운 유형 불량도 척척 잡아... AI 만난 스마트팩토리 진화", 동아일보, 2023. 06. 19.
- [2] C. Lugaresi, et al. "Mediapipe: A framework for building perception pipelines.", 2019.
- [3] 임대환, 남기훈. "한글 단어 및 문장의 음성 인식을 향상을 위한 음성 인식 API 개발", 한국정보기술학회논문지, 2019, pp. 95-104.

제품이미지를 활용한 웹드라마 광고콘텐츠 제작사례

이채이, 이채영, 은광하
 상명대학교 디지털콘텐츠전공

e-mail : dmgh12@naver.com, color3810@gmail.com, kheun0310@smu.ac.kr

A Case of Production of Web Drama Advertising Content Using Product Image

Chaeyi Lee , ChaeYoung Lee , Eun, Kwang Ha
 Major in Digital Contents, Sangmyung University

요 약

본 논문은 기업 홍보용 광고 제작사례를 제시한다. 작품에서 활용한 제품은 기업 ‘c’사의 제품으로 ‘c사’ 브랜드 인지도 제고를 의도로 기획되었다. 1인 가구가 늘어나며 밥솥시장이 축소되고 있는 상황에서 밥솥의 가치를 증명하기 위해 감성마케팅을 활용 하였으며 설정한 캐릭터와 등장인물의 관계성, 스토리텔링을 통해 신뢰감을 주는 브랜드 이미지를 위한 플롯으로 구성하였다. 또한, 대본을 통해 제품의 다양한 기능과 실질적인 활용방안을 직접적으로 제시 함으로 홍보용 콘텐츠를 제작하였다.

1. 서론

웹드라마를 통한 제품홍보 영상은 저변에 확대되어 다양한 서비스의 제품을 소개하는 광고형태의 웹드라마가 제작되고 있다. 이번 논문은 브랜드 ‘C사’의 마케팅 전략으로 메이저 광고기획 아이디어 페스타를 통해 대학생들에게 광고영상 제작가이드를 기반으로 웹드라마 기반 제작 사례를 제시하고자 한다. 주요 제시 기준은 핵심타겟을 20대 후반~30대 후반의 M세대로 경쟁사와는 차별화된 ‘C사’의 밥솥 브랜드 아이덴티티를 활용하여 ‘C사’의 브랜드 자산을 증대 시킬 수 있는 전략을 모색함과 M세대 대상 프리미엄 가격의 지불, 긍정적 구전, 높은 인게이지먼트 달성 등 ‘브랜드 공명’에 도달할 수 있는 IMC 아이디어를 제시하도록 하였다. 또한 현재 여러 난제를 극복하기 위한 ‘사’의 밥솥 마케팅 솔루션 제안 또는 M세대가 공감 가능한 다양한 톨의 영상 콘텐츠를 제작하는 것이다. 현재 밥솥시장은 1인가구의 증가와 간편식, 배달문화의 확산으로 밥솥시장이 급격히 축소되고 있는 상황이었다. C사’의 밥솥은 양극화 체제 시장에서 브랜드 파워를 지켜내야 하지만 기업의 프리미엄 브랜드 이미지를 고수하기에 경쟁사의 제품과 기능적 측면에서 차별점이 없다는 것이 당면한 문제해결 상황이었다. 따라서 제작 방향은 제품의 차별이 어려운 상황 속 제품모던 이미지차별을 시도하기로 하였고 브랜드 이미지 광고를 기획하게 되었다. 현재 ‘C사’의 밥솥의 메인 모델 ‘B 밥솥’은 밥솥 자체를 의인화한 전략으로 홍보되고 있었다. 따라서 기존의 제품 ‘B 밥솥’으로 ‘C사’의 브랜드 이미지를 창출할 수 있는 디지털 마케팅 전략을 연결하여 ‘B 밥솥’ 광고를 의인화 방향으로 제작하고 ‘C사’의 공식 유튜브채널에 업로드하여 시청자의 반응을 확인하고자 하였다.

2. 본론

2.1 기획배경

현재 유튜브 시장에서 흥행하고 있는 광고콘텐츠의 사례를 분석해보았을 때 라네즈 래디언씨 크림 광고나 HD현대의 내 땅강아지, 대밭이 광고 등 높은 퀄리티의 영화적 접근과 신선한 연출의 집점으로 거부감 없이 광고효과를 내고 있었다. 이러한 영화적 연출이 ‘반려 밥솥’ 키워드로 스토리텔링을 보여주는데 적합하다고 판단하였다. 기획된 작품도 마찬가지로 영화적 연출을 활용하였으며 스토리 전개 속에서 ‘B 밥솥’과 등장인물의 관계를 통해 소비자와 기업의 관계를 상징적으로 연출하였다.

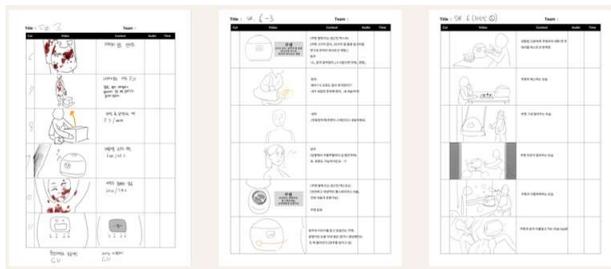
(표 1) 웹드라마 제작 개요

제목	세상에 나쁜 밥솥은 없다.
로그라인	간편식에 밀려 소외된 밥솥 ‘쿠첸’이 가출하게 되면서 벌어지는 성장스토리
장르	단편 실사 웹드라마(가족 코미디)
홍보 매체	쿠첸 공식 유튜브
소재	‘C 사의 전기압력밥솥
주제	밥솥의 가능성과 소중함
기획목적	브랜드 인지도 제고, 구매촉진, 정보전달

2.2 웹드라마 제작 내용

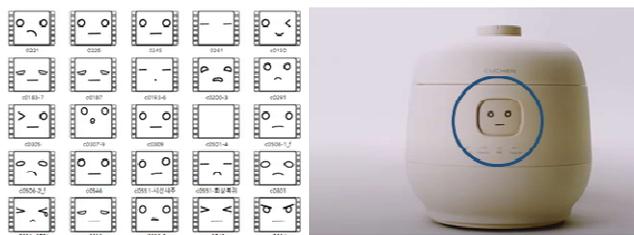
1인 가구비율의 증가와 즉석밥, 간편식, 배달문화의 발달

로 실생활에서 밥솥의 사용빈도가 줄어든 것과 함께 밥을 짓는 행위 자체가 줄어들며 밥솥의 필수 가전으로서 입지가 축소되었다는 점에서 문제점을 규명하였다. 밥을 밥솥으로 직접 지어먹을 필요가 없어진 상황 속 이러한 문제점을 타개하기 위해서 우린 압력밥솥이 밥뿐만 아니라 다른 요리에도 유용하게 쓰일 수 있다는 점을 스토리텔링을 통해 보여 줘야만 했다. 제작된 웹드라마의 핵심 메시지는 ‘다 잘하는 기특한 쿠첸’으로 반려 밥솥 브레인의 이야기를 담은 웹드라마 시리즈 3편과 밥솥으로 만들 수 있는 요리 영상 1편 총 4편의 영상로 구성되어 있다. 기존의 cf에서 볼 수 있는 ‘B 밥솥’의 기본표정에서 더 나아가 스토리텔링을 담을 수 있는 다양한 표정을 제작하였다. 본 영상 촬영을 위해서 스토리보드를 [그림 1]과 같이 작성하여 장면 단위 점검 지점을 확인하였다.

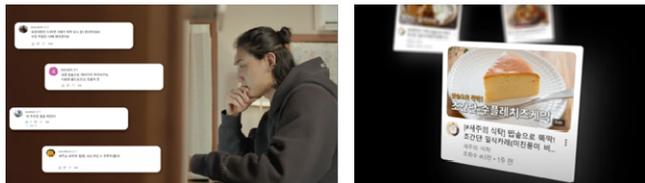


(그림 1) 촬영 스토리보드

선별된 표정들은 애니메이션으로 제작한 뒤 밥솥에 합성하는 트래킹 작업을 하였다. 또한, 그래픽 파트에서 요리 유튜버 새주가 복분자 식혜 영상을 올렸을 때의 반응을 극적으로 표현하여 밥솥이 밥 만 이 아닌 다양한 레시피로 활용될 수 있다는 점을 효과적으로 인식시켰다.



(그림 2) 캐릭터 표정 합성

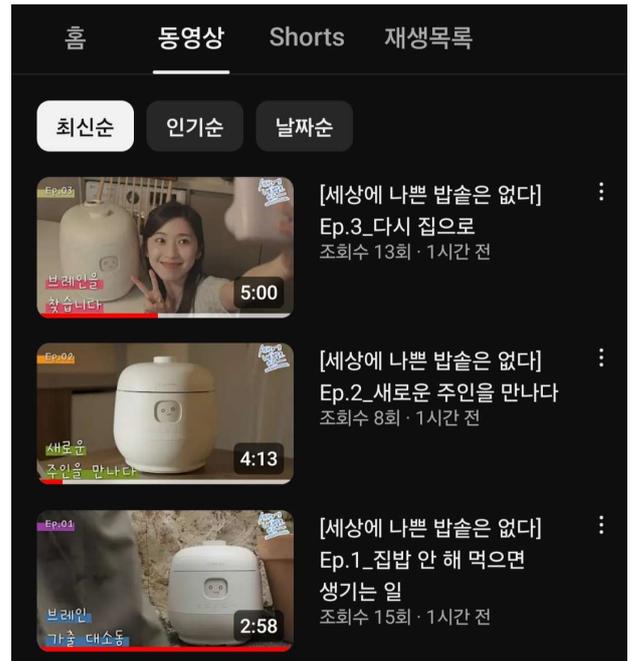


(그림 3) 웹드라마 장면

3. 기대효과

기획된 내용을 바탕으로 제작된 작품 ‘세상에 나쁜 밥솥

은 없다.’는 ‘C사’의 ‘B 밥솥’을 의인화 전략으로 시청자에게 브랜드 애착을 형성하게 한다. 이를 통해 브랜드 인지도를 높이고 브랜드 공명을 의도할 수 있으며 압력밥솥의 다양한 활용성을 인식시켜 타겟층의 구매를 촉진하는 매개체가 될 수 있다. 또한, 콘텐츠 포맷의 특성을 이용해 제품 정보를 거부감 없이 전달할 수 있는 장점이 있다. 제작된 콘텐츠는 ‘C사’와 협의를 통해 해당 회사의 공식 유튜브 채널에 업로드되었으며 이는 학생과 기업이 연계하여 브랜드 홍보 콘텐츠를 제작하였다는 데서 또 한 번의 의의를 가진다.



(그림 3) 유튜브 채널에 업로드 된 웹드라마 광고

참고 문헌

- [1] 강신용. "국내 장수브랜드의 새로운 광고표현 제작사례", 『일러스트레이션 포럼 Vol.59』, p6
- [2] 강신용. "제품브랜드를 활용한 다문화 광고 제작사례", 『일러스트레이션 포럼 Vol.55』, p57
- [3] 이홍욱. "유튜브광고 콘텐츠속성이 제품구매의도에 미치는 영향 : 브랜드인지도와 제품이미지의 매개효과를 중심으로", 『식사학위 프로젝트 논문』, p27

프로젝션 맵핑 기반 고전 게임의 표현 적용을 위한 시각표현 연구

김동민, 박수민, 이규서, 은광하
 상명대학교 디지털콘텐츠 전공

e-mail : ohohdongmin@naver.com, soomin141108@naver.com,
 jusarang10104@naver.com, kheun0310@smu.ac.kr

Visual representation research for projection mapping-based representation of classic games

Kim dong-min, Park soo-min, Lee kyu-seo, Eun kwang-ha
 Major in Digital Contents, Sangmyung University

요약

본 연구는 프로젝션 맵핑 기술을 활용하여 건물 외벽 구조물에 향수를 불러일으킬 수 있는 레트로 주제를 선정하여 효과적인 표현을 위하여 시각표현 연구를 통해서 3가지 제작 키워드를 중심으로 고전 게임의 플레이 형태를 분석하여 미디어아트 콘텐츠를 제작하였다. 이를 통해서 완성된 프로젝션 맵핑 콘텐츠의 활용방안을 제시하였다.

1. 서론

미디어아트는 최근 다양한 디스플레이 기술과 연동되어 대중들에게 선보이고 있다. 현대 시대의 과학 기술과 문화를 이용하여, 게임이나 영상 등 다양한 매체로, AR, VR 등 다양한 기술과 결합하여 더욱 관객과 상호작용하는 방식으로 표현되고 있다. 본 연구는 2000년대 이전의 고전 게임, 속칭 ‘레트로(Retro)’를 주제로 빔프로젝터를 이용하여 구조물에 투과시키는 방식의 프로젝션 맵핑 콘텐츠의 제작 및 전시를 통해서 이용자의 반응을 살펴보고 해당 주제의 시각표현에 대한 제작 사례를 언급하고자 한다.

2. 본론

2.1 프로젝션 맵핑 기반 영상 콘텐츠 제작 과정

1)주제 선정

프로젝션 맵핑 콘텐츠의 주제 선정 과정으로, 고전 게임 주제 작품 조사를 진행 하였고, 기존 미디어 아트는 그림 1과 같이 벽에 빔을 조사, 전시를 진행함을 확인하였다.



(그림 1) 게임 슈퍼마리오 게임 주제 미디어아트

그림 1과 같이 인기 IP인 ‘슈퍼 마리오’를 주제로 한 게임의 미디어 아트를 참고, 프로젝션 맵핑 콘텐츠에 적용시키기 위하여 과거 유행한 레트로 게임의 조사를 진행하였다. 그림 2와 같이 3종의 게임 ‘동키콩’, ‘젤다의 전설’, ‘배틀 시티(탱크게임)’를 선정하였으며, 해당 게임들을 재해석 하여 고양이가 게임 세계를 모험한다는 주제의 프로젝션 맵핑 콘텐츠를 제작하였다.



(그림 2) 동키콩(좌), 젤다의 전설(중앙), 배틀 시티(오른쪽)

2) 프로젝션 맵핑 미디어 아트 제작 과정

선정한 주제 기반의 미디어 아트는 레트로 게임의 재해석을 목표로 하여, 고전 게임의 요소인 고대비, 도트 그래픽을 비롯한 픽셀 아트, 레트로를 표현할 수 있는 요소 추가라는 3개의 주요 키워드로 삼아 제작에 참고하였다. 그림 3과 같이, 고대비의 색상을 사용하여 보다 원근감의 구분이나, 캐릭터와 배경 요소의 구분감을 극대화 하였다. 또한 레트로 표현을 위하여 작품에 고전 게임기의 형태의 프레임을 삽입하기로 결정하였으며, 추가적인 구현을 위해 조이스틱, 4종의 버튼, 동전 삽입구를 추가하였다. 또한, 픽셀 아트 구현을 위해 “Aseprite” 프로그램을 통해 픽셀 그래픽을 구현하였고, 이후 완성된 미디어 아트를 본교 종합실기관의 격자무늬 구조물을 구현한 3D파일에 맞추어 기울기 등을 조절 및 변형하여 “Adobe After Effect” 프로그램에 삽입하여 영상화 프로세스를 완료하였다.

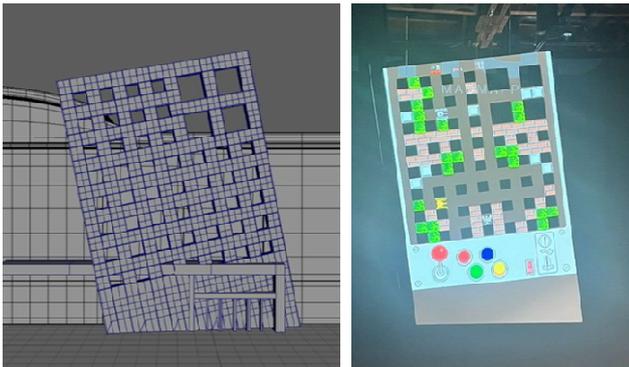


(그림 3) 프로젝션 맵핑 미디어 아트 구현

2.2 미디어 아트 전시 과정

1) 프로젝션 맵핑 미디어 아트 전시 과정

제작된 미디어 아트는 본교 종합실기관의 격자무늬 구조물을 기반으로 하여 전시를 목표로 하였다.



(그림 4) 구조물에 적용한 프로젝션 맵핑 영상 콘텐츠

제작 완료된 미디어 아트의 전시를 위하여 그림 4와 같이 격자무늬 구조물과 빔프로젝터를 전시장에 설치하였으며, 프로젝션 맵핑을 위하여 'MadMapper' 프로그램을 사용하여 구조물에 맞춘 맵핑 작업을 완료하였다.

2) 프로젝션 맵핑 미디어 아트 전시 반응 분석 과정

전시 동안 대면 인터뷰 형식으로 표 1과 같이 3종의 문항을 통해 10 ~ 20대 전시 관람객을 대상으로 전시된 미디어 아트 작품의 반응을 집계하였다.

(표 1) 전시 작품의 대면 인터뷰 항목

인터뷰 항목	문항
문항 1	전시 작품의 주제를 몰랐을 때 감상은 어땠습니까?
문항 2	전시 작품의 주제를 알았을 때 감상은 어땠습니까?
문항 3	전시 작품의 아쉬운 점은 무엇이라 생각하십니까?

3. 결 론

3.1 게임 기반 프로젝션 맵핑 미디어 아트의 시각 표현 분석

본 연구에서 제작된 고전 게임을 기반으로한 프로젝션 맵핑 미디어 아트 작품에 사용된 시각 표현으로는 픽셀 그래픽, 고대비, 레트로 표현 3종이 있다. 이 중, 픽셀 그래픽과 고대비 2종의 시각 표현은 레퍼런스로 선정한 3개의 게임에 공통적으로 나타나는 표현으로, 고전 게임들의 제작 시기인 1980년대~1990년대 당대의 게임기기의 용량적 한계로 인한 픽셀 그래픽의 사용, 디스플레이 기술 한계로 인한 고대비를 통한 고시인성을, 2종의 표현을 제외한 레트로 표현은 현대의 디지털 기술과 차별감을 주기 위한, 사용되지 않는 공간에 재미와 주제의 통일감 형성을 위한 고전 게임기 형태의 프레임을 적용함으로 3종의 표현의 조화를 통해 레트로 문화에 대한 향수와 공감대를 유도할 목적으로 제작됨을 알 수 있다.

3.2 프로젝션 맵핑 미디어 아트의 전시 반응 분석

본 연구에서 제작된 프로젝션 맵핑 활용 전시 작품의 반응 분석 결과, 공통적으로 표 2의 반응으로 집계되었다.

(표 2) 전시 작품의 대면 인터뷰 설문 반응 분석

인터뷰	반응
문항 1	레트로 게임을 추억하는 느낌이었다.
문항 2	기성 세대와 이야기할 수 있는 공통 주제가 생긴 것 같다.
문항 3	개별 게임에 대한 제반 지식이 부족하면 이해가 어려울 수 있을 것 같다.

3.3 프로젝션 맵핑 미디어 아트의 기대 효과

본 연구에서는 프로젝션 맵핑을 활용한 고전 게임 기반의 미디어 아트 제작 사례를 바탕으로 표현 적용을 위한 시각적인 연구를 진행하였다. 프로젝션 맵핑 방식의 활용의 확대를 위하여 현재 유행하고 있는 레트로 문화를 주제로 다양한 표현기법을 사용하여 건물의 외벽을 비롯, 다양한 구조물에 적용할 수 있는 방안을 전시 종료 이후, 추가적인 연구를 통해서 제시하고자 한다. 이는 단순 시각적으로 관찰되는 기존의 미디어아트와는 달리 레트로 문화를 프로젝션 맵핑 기술과 결합해 적용하여, 기성 세대에게는 향수와 문화적 공감대를 자극하고, 10대와 20대를 비롯한 젊은 세대에는 새롭고 몰입감 있는 문화적 경험과 기성세대와 함께 느낄 수 있는 문화적 공감대를 형성하는 효과를 기대할 수 있다 생각함과 동시에, 기존의 대중문화와 결합하여 새로운 경험을 창조할 수 있음을 기대한다.

참 고 문 헌

- [1] 김경은. “홀로그램 실감 콘텐츠를 통한 미디어 아트 시각 표현 연구 -본인의 작품을 중심으로-”, 『예술과 미디어』 제21권 제1호, 2022. pp.31-54.
- [2] 김희숙. “대중을 위한 소통의 플랫폼으로서 미디어파사드의 활용성에 관한 연구”, 『Journal of Cultural Product & Design』 Vol. 74, 2023. pp.319-328.

시각장애인의 게임 접근성 향상을 위한 핸드트래킹 게임 개발

*곽수연, **김현진, **김연아, *황준, ***고진광
*서울여자대학교 소프트웨어융합학과, **서울여자대학교 디지털미디어학과,
***순천대학교
e-mail : kimyeona0926@gmail.com

Hand Tracking Game to Improve Game Accessibility for The Visually Impaired

*Suyeon Kwak, **Hyunjin Kim, **Yeona Kim, *Jun Hwang, ***Jingwang Go
*Department of Software Convergence, Seoul Women's University,
**Department Digital Media Design and Applications, Seoul Women's
University, ***Suncheon National University

요 약

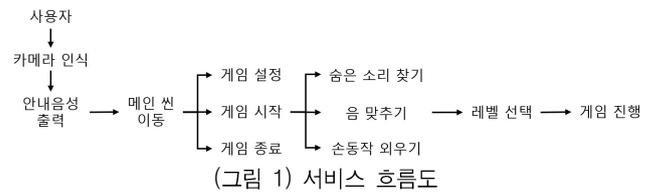
시각장애인과 비시각장애인의 동일한 게임 경험을 목표로, OpenCV와 CVZone 라이브러리, 손동작 인식 인공지능과 유니티를 활용하여 핸드 트래킹 기술과 음성 사용자 인터페이스로 구성된 게임을 개발하였다. 소리를 이용한 3가지 게임들로 게임마다 난이도를 설정할 수 있도록 하여 게임의 재미성을 높였다. 이를 통해 시각장애인과 비시각장애인의 문화 격차 해소, 공감대 형성에 도움을 제공할 것으로 기대한다.

1. 서 론

오늘날 디지털 사회에서 디지털 접근성이 강조되면서 시각 장애의 여부에 영향을 받지 않도록 하는 게임 접근성 개선에 대한 필요성이 대두되고 있다. 그러나 여전히 시각 장애인들이 게임을 진행하는 데에 여러 제약들이 존재한다. 시각장애인들은 게임을 이용할 때 불편을 해결하기 위한 별도의 보조기기 프로그램 및 소프트웨어에 가장 많이 의존할 수밖에 없으며[1], 마우스 사용의 어려움으로 기타 다른 물리 입력 장치를 사용해야 한다. 또한, 음성출력 및 화면 확대와 같은 게임 자체 수준에서의 접근성 기능 탑재가 요구되고 있으나 현재 수준에서는 지원되지 않는 경우가 많아 시각장애인은 별도의 외부적인 접근 방법에 의존할 수밖에 없다.[1, 2] 더 나아가 이러한 외부적인 접근 방법이 오히려 게임을 이용하는 데 있어 도움이 되지 않는 경우도 있기 때문에 접근성 개선 방안을 고안하는 데에 더욱 신중을 가해야 한다. 시각장애인이 외부적인 접근 방법으로 게임을 즐기는 사례로서 [3]의 경우 시각장애인을 위한 모바일 기기의 스크린리더 기능을 제공한다. 그런데 해당 정보를 음성 메시지로 명확히 전달하지 않아 원활한 게임 접근이 불가하다는 한계가 있다.[1]

본 연구에서는 음성 안내와 같은 게임 자체의 접근성 기능을 갖추고 단순한 손동작을 게임의 주요 조작 방법으로 하는 게임을 제안한다. 이를 통해 별도의 보조기기 프로그램 및 소프트웨어를 사용하지 않아도 시각장애인과 비시각장애인이 가능한 동일한 게임 경험을 가질 수 있도록 한다.

2. 연구 내용



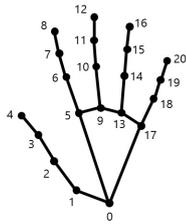
본 게임은 그림 1과 같이 동작한다. 본 연구의 플랫폼으로는 직관적인 개발 인터페이스와 게임 제작에 필요한 다양한 기능을 제공하는 유니티를 선택하였으며, 이를 기반으로 여러 기능 및 게임들을 제작하였다.

손 인식에 필요한 관련 기술은 OpenCV 등 여러 Open API를 활용하여 구현하였다. 게임 전체에서 자주 사용되는 요소를 효율적인 방식으로 구현하기 위해 핸드트래킹 씬은 Additive Scene으로 제작하였다. 해당 방법을 통해 기능 및 역할별로 씬을 분리해 주어 리소스를 효율적으로 관리하고 협업 과정에서 충돌을 줄여 작업을 용이하게 하였다. 또한 씬 전환 기능은 싱글톤으로 작성하여 원하는 스크립트 어디에서든 쉽게 접근할 수 있도록 하였다.

그 외에도 시스템의 확장성과 유연성을 높이기 위해 이벤트 기반의 구조를 택하였다. 또한, 오브젝트들이 ScriptableObject 를 통해 이벤트를 호출하고 리스너하도록 하여 객체간의 의존성을 낮추어 생산성을 높이며 코드의 가독성을 개선하였다.

게임의 접근성을 높이기 위해 키보드로도 메뉴 선택 및 게임 안내 음성 듣기를 할 수 있도록 하였다.

2.1 손 인식 구현



(그림 2) 손의 랜드마크

손 인식은 손의 21개의 랜드마크 좌표를 통해 이루어진다. 실시간으로 손 위치와 손동작을 인식하기 위해 OpenCV, CVZone 라이브러리, 손동작 인식 인공지능 모델을 사용하여 1초에 10번 웹캠에 인식되는 손의 랜드마크 좌표들과 손동작 인식 결과를 가져오도록 하였다.

카메라에 게임에서 사용되는 손 이외의 손이 인식되는 것을 방지하기 위해 최대 두 개의 손이 인식되도록 하였다. 두 개의 손이 인식되었을 경우 그림 2의 랜드마크 5번과 17번 좌표의 거리와 그때의 실제 카메라에서 손까지의 거리 사이의 관계를 나타내는 이차함수를 통해 두 손의 카메라로부터의 거리를 구한다. 두 손 중 거리가 더 가까운 손을 게임에 사용되는 손으로 인식하고 해당 손의 거리가 카메라로부터 너무 멀거나 가깝지 않은지, 현재 손동작이 무엇인지 판단한다. 마지막은 손 랜드마크 좌표, 손의 거리 판단 결과, 손동작 인식 결과를 UDP로 서버에 보낸다. 카메라에 손이 벗어날 수 경우에는 손의 정보가 아닌 "out"이라는 문자를 UDP로 서버에 보낸다.



(그림 3) 카메라 캡처 이미지



(그림 4) 전처리 후 이미지

게임에서 사용되는 손동작은 그림 3과 같이 세 가지이다. 보자기는 기본적인 손동작으로 게임의 오브젝트를 감지한다. 주먹은 선택 손동작으로 게임 내 오브젝트를 선택한다. 가위는 게임 설명을 듣기 위한 손동작이다.

손동작을 구분하기 위한 인공지능 모델은 Google Teachable Machine 사용하여 제작하였다. 손이 인식된 카메라 화면을 캡처한 이미지를 그림 4와 같이 300x300 크기의 이미지로 전처리하여 모델 학습에 사용하였으며 각 손동작 당 약 4,000장씩 학습시켰다. 모델의 파라미터로 에포크 100, 배치 크기 512, 학습률 0.00105로 설정하였다. 약 4,000장의 이미지 중 약 600장을 테스트로 사용하였다.

손동작을 학습시킨 결과 Google Teachable Machine에서 모델의 성능으로 0.99의 테스트 정확도와 약 0.001의 테스트 손실을 보였다. 손동작에 대한 정확도는 세 가지

손동작 모두 1을 보였다. 그러나 실제로 게임 내 적용하였을 때 손이 카메라로부터 너무 가깝거나 멀 경우 손동작을 제대로 구별하지 못하는 모습을 보였다. 이는 손의 랜드마크 거리가 너무 가깝거나 멀어 발생한 문제로 추정되었다. 따라서 해당 거리에 제한을 두어 손이 제한 거리를 벗어나지 못하도록 게임 내에서 안내하였다.

2.2 미니게임 제작

게임성을 높이고 사용자에게 다양한 경험을 제공하기 위하여 표 1과 같이 세 가지 미니게임을 구현하였으며, 각 미니게임 별로 상, 중, 하의 세 가지 난이도를 제공하였다.

게임의 청각적 요소의 비중이 높은 만큼 사용자가 반복해서 들었을 때 피로감을 느끼지 않는 소리 리소스를 적용하였다.

저시력 시각장애인과 색각이상자의 접근성 향상을 위해 게임 화면을 구성하는 색상들을 고대비 조합으로 구성하여 인지성, 명확성을 높여주었다.

(표 1) 구현된 미니게임 종류 및 설명

미니게임	게임 설명
숨은 소리 찾기	격자로 나열된 블록 중 좌우에서 다른 소리가 재생되는 블록을 찾는 게임
음 맞추기	손을 위아래로 움직여 음을 조절해 주어진 음과 동일하게 맞추는 게임
손동작 외우기	음성으로 들려주는 손동작을 기억하는 게임

3. 결 론

본 논문에서는 별도의 보조기기 및 프로그램을 사용하지 않는 핸드트래킹과 음성 중심의 인터페이스를 제공하는 게임을 제안하였다. 시각장애인의 게임 접근성 개선에 핸드트래킹 기술이 도움이 될 수 있는지 확인할 수 있었으며, 이를 통해 시각장애인과 비시각장애인의 문화 격차 해소, 공감대 형성에 도움을 제공할 것으로 기대한다. 앞으로 게임 자체 수준에서의 접근성 기능이 탑재된 게임이 많이 만들어지기를 바란다.

ACKNOWLEDGEMENT

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)

참고 문헌

[1] 한국콘텐츠진흥원, "시각장애인 게임 접근성 개선 방안 연구", 2024. pp.106-108 pp.109 pp.144-148 pp.150-151.
 [2] 한국콘텐츠진흥원, "장애인 게임 접근성 제고 방안 기초 연구", 2023. pp.84-86 pp.109-111 pp.130.
 [3] NEXON Company. 메이플스토리 M, 2023. Google Play Store

졸업 프로젝트 아카이빙과 신입 인재 채용을 위한 양방향 매칭 플랫폼에 대한 연구

김주은, 김희주, 황준
 서울여자대학교 소프트웨어융합학과
 e-mail : jueunk617@swu.ac.kr

A Study on an Interactive Platform for Matching Graduation Projects with New Talent Recruitment

Jueun Kim, Huiju Kim, Jun Hwang
 Department of Software Convergence, Seoul Women's University

요 약

최근 졸업 예정자의 취업이 어려워지면서, 졸업 프로젝트의 체계적 관리와 이를 활용한 신입 인재 채용의 중요성이 대두되고 있다. 본 논문은 다양한 분야의 졸업 프로젝트를 아카이빙하고 신입 인재를 효과적으로 발굴할 수 있는 양방향 통합 플랫폼의 필요성과 구현 방안을 제시한다. 이를 통해 학생, 대학, 기업 간의 협력을 강화하고, 인재 선발의 효율성을 높이는 데 기여하고자 한다.

1. 서 론

본 논문은 졸업 프로젝트 아카이빙과 신입 인재 채용을 효과적으로 지원하는 통합 플랫폼의 개발을 목표로 한다. 졸업 프로젝트는 대학에서 배운 지식과 기술을 활용하여 학부생의 학문적 종합 산출물을 만든다는 의미에서 능력을 보여주는 데 큰 역할을 한다. 그러나 대학의 졸업 프로젝트는 장기적으로 관리되지 않아 학문적 자산으로서 제대로 활용하지 못하고 있다. 또한, 기업은 원하는 인재를 찾지 못하고 학생은 역량을 입증하기 힘들어하는 일자리 미스매치 현상[1]이 나타나고 있다.

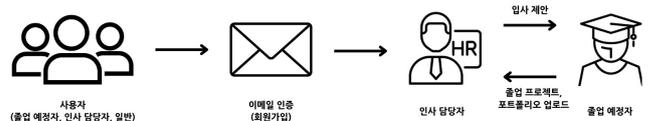
현재 앱 시장에서 사용되는 채용 플랫폼들은 경력 정보와 인적 네트워크에 중점을 두지만, 졸업 프로젝트와 커리어를 효과적으로 연계하는 기능은 부족하다. 대표적으로 [2]와 [3]은 이력서 중심의 경력 관리 시스템이기 때문에 연구 성과나 프로젝트 기반의 포트폴리오를 강조하기 어렵고, [4]는 주로 신입보다는 경력직을 위한 네트워크에 초점이 맞춰져 있기 때문에 졸업 프로젝트와 같은 학생의 성과를 보여주기 어렵다. 따라서 학생과 기업이 졸업 프로젝트를 체계적으로 관리하고, 이를 신입 인재 채용에 활용할 양방향 통합 플랫폼이 필요하다.

본 연구에서는 효과적인 인재 채용을 위한 졸업 프로젝트 아카이빙 및 채용을 위한 양방향 플랫폼을 제안한다. 이 플랫폼은 기존 채용 플랫폼의 한계를 넘어 졸업 프로젝트와 같은 학문적 산출물을 중심으로 신입 인재의 능력을 보여줄 수 있다는 점에서 차별성을 가진다. 기업 역시 단순 경력이 아닌 프로젝트 중심의 능력 검증이 가능하다. 또한, 채용과 프로젝트 게시가 동시에 이루어지기 때문에 기존 플랫폼과 다른 새로운 채용 방식으로 신입 인재와 기업 간 빠르고 정확한 채용 환경을 조성하고 양방향 소통을 활성화할 수 있다.

2. 연구 내용

2.1 서비스 개요

본 플랫폼의 서비스 주요 흐름도는 그림1과 같다. 신뢰도 있는 앱 환경을 위해 사용자의 신원 인증 절차를 거치는데, 이는 사용자 유형에 따라 다르게 구성된다. 졸업 예정자는 학교 이메일 인증 및 재학생증서를 제출하고, 인사 담당자는 회사 이메일 인증과 재직증명서를 제출해야 한다. 일반 사용자는 별도의 인증 절차 없이 로그인할 수 있다. 그러나 인증을 해야만 피드 및 공고 제안서 작성과 같은 양방향 소통이 가능하다. 사용자 전체는 업로드된 프로젝트를 피드로 열람하고, 플랫폼의 자체 제작 콘텐츠를 확인한다. 추가로 해시태그 검색과 학교/학과 검색 기능을 활용하여 원하는 조건의 피드를 볼 수 있다.



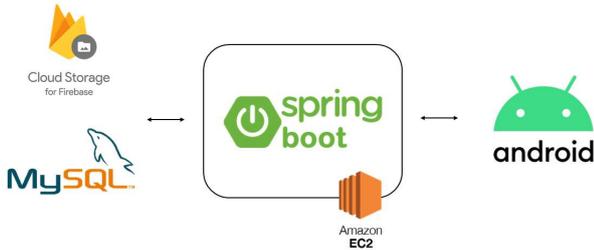
(그림 1) 서비스 주요 흐름도

졸업 예정자는 프로젝트 업로드가 가능하고, 해당 프로젝트에 참여한 팀원을 초대할 수 있다. 졸업 프로젝트의 진시회가 열리는 경우, 진시 장소와 일정도 작성이 가능하다. 또한, 인사 담당자는 인상 깊은 프로젝트나 프로젝트에 참여한 학생을 스크랩할 수 있다. 스크랩된 학생의 역할과 기업의 직무가 적합하다면 제안서를 보내 채용 의사를 학생에게 전달할 수 있다. 프로젝트 아카이빙 및 채용 과정이 이러한 절차로 동시에 진행된다.

2.2 서비스 구현 환경

본 연구의 서비스 구현 환경은 그림2와 같다. 서버는 Spring Boot로 구현되었으며, AWS EC2 인스턴스에서 호

스팅 되어 높은 신뢰성과 확장성을 제공한다. DBMS는 MySQL을 활용하여 데이터를 효과적으로 처리한다.



(그림 2) 서비스 구현 환경

서버는 REST API를 통해 클라이언트와 통신한다. Entity, DTO, Controller, Service, Repository 구조로 계층적 아키텍처를 구현한다. 사용자 인증은 JWT를 이용하여 처리되며, Access Token과 Refresh Token이 사용된다. 클라이언트 측에서는 MVP 패턴의 요소를 부분적으로 적용하여 Android 애플리케이션으로 구현되었다. 사용자가 서버에 요청하고 응답을 받아서 사용자 인터페이스에 직관적인 컴포넌트들이 표시된다. 학교, 학과 검색 기능은 커리어넷의 OpenAPI를 활용하여 구현되었다. 이미지는 multipart/form-data 형식으로 Firebase Storage에 저장된 후 연동하여 체계적으로 관리한다.

2.3 서비스 주요 기능

(표 1) 서비스 주요 기능 설명

주요 기능	기능 설명
해시태그 기반 프로젝트 업로드 및 검색	해시태그를 통한 세분화된 프로젝트 검색 제공
학교/학과 피드 설정 및 검색	학교/학과별 프로젝트 검색 및 피드 설정
졸업 예정자 프로필 포트폴리오	다양한 성과물 업로드 및 포트폴리오 관리
인증된 인사 담당자의 채용 제안	인증된 인사 담당자만 입사 제안 가능
프로젝트 큐레이팅 콘텐츠: Stage	주목할 만한 프로젝트 및 졸업 예정자 심층 소개
관심 있는 게시물, 콘텐츠, 인재 스크랩	관심 게시물, 콘텐츠 인재 스크랩 및 저장 관리

졸업 프로젝트 아카이빙의 핵심은 다양한 프로젝트들을 어떻게 분류하고 검색할 수 있는지에 있다. 홈 피드의 분야별 카테고리만으로는 자세한 검색과 인재 탐색에 한계가 존재한다. 따라서 해시태그를 통해 더욱 세분된 정보와 키워드로 사용자의 관심사에 따른 프로젝트 업로드와 검색 기능을 제공한다. 사용자가 졸업 프로젝트를 탐색하는 과정에서는 분야뿐만 아니라 특정 학교, 학과에 대한 궁금증도 발생한다. 타 학교의 작품도 조회가 가능하도록 학교/학과 검색 기능을 제공하였다. 프로젝트만으로는 제공되는 정보가 한정적이기에 학생은 포트폴리오를 통해 다양한 성과물과 경험을 체계적으로 관리한다. 인사 담당자는

이를 조회하여 전반적인 역량을 확인하고 종합적으로 졸업 예정자에게 적합한 직무를 제안할 수 있다. 모집 직무, 연락처를 포함한 상세한 입사 제안서가 전달되며, 이를 통해 명확한 정보 제공과 원활한 양방향 소통이 이루어진다. ‘Stage’는 프로젝트 큐레이팅 기능을 갖춘 콘텐츠로, 주목할 만한 프로젝트와 졸업 예정자를 심층적으로 소개한다. 이 화면은 각 프로젝트와 졸업 예정자의 인터뷰를 포함한 상세 정보를 제공하며, 하단에는 해당 프로젝트 링크와 전시 정보도 기재하여 추가적인 홍보를 지원한다. 모든 사용자는 스크랩을 통해 관심 있는 졸업 프로젝트 게시물과 큐레이팅 콘텐츠를 저장한다. 추가로 인사 담당자는 관심 있는 인재까지 스크랩할 수 있다. 이 기능을 통해 사용자는 자주 참조하는 자료를 체계적으로 정리하고, 인사 담당자는 선별된 인재를 효율적으로 추적하며 관리할 수 있다.

2.4 결론

본 연구에서는 졸업 예정자의 졸업 프로젝트 아카이빙과 인사 담당자의 신입 인재 채용을 지원하는 양방향 플랫폼을 개발하였다. 학생들은 프로젝트 방향 설정에 필요한 구체적인 자료와 사례를 쉽게 확보하여, 졸업 프로젝트의 질적 향상을 이끌어 낼 수 있다. 졸업 예정자들이 프로젝트를 게시해 인사 담당자들에게 역량을 효과적으로 입증할 수 있고, 기업은 이를 바탕으로 잠재적인 신입 인재를 발굴하고 인재 채용의 효율성을 높일 수 있다. 이러한 서비스 과정에서 네트워킹의 장이 형성되어 사용자 간 지식과 정보의 교류가 활성화되고 다양한 협업의 기회가 창출된다. 인사 담당자는 최신 졸업 프로젝트를 통해 현재 학생들의 관심사를 파악할 수 있으며, 이는 기업의 연구개발 및 혁신에 긍정적 영향을 미칠 수 있다.

본 플랫폼은 다국어 지원 및 번역 기능을 추가하여 세계 시장으로의 확장이 가능하다. 교육과 취업의 경계를 넘어 국제적 차원에서 혁신적인 학습 및 채용 패러다임이 구축될 것으로 기대된다.

ACKNOWLEDGEMENT

“본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)”

참고 문헌

[1] 2024년 상반기 대기업 채용동향인식 조사(2024). https://www.fki.or.kr/main/news/statement_detail.do?bbs_id=00035481&category=ST, (9, 5, 2024).
 [2] 잡코리아. (2012). 잡코리아 앱(버전 4.0.22) [모바일 앱]. Google Play Store
 [3] (주)사람인. (2014). 사람인 (버전 3.5.7) [모바일 앱]. Google Play Store
 [4] LinkedIn. (2011). LinkedIn 앱(버전 4.1.979) [모바일 앱]. Google Play Store

AI 기반 반려묘 피부 증상 탐지 서비스

최소원*, 박유진**, 이예림**, 엄성용*
 서울여자대학교 소프트웨어융합학과*, 서울여자대학교 디지털미디어학과**
 e-mail : {weileenw, youjin09222, yearim2008, osy}@swu.ac.kr

AI-based cat skin symptom detection

So Won Choi*, You Jin Park**, Ye Rim Lee**, Seong Yong Ohm*
 Seoul Women's University Department of Software Convergence*,
 Seoul Women's University Department of Digital Media**

요 약

반려묘 피부 증상 탐지 서비스로, 고양이의 피부 질환이 의심되는 부위의 사진을 4장 이상 찍어 제출하면 의심 증상 부위에 대한 정보와 함께 의심 증상에 대한 정보를 제공하는 서비스이다. 이미지 내에서의 피부 병변을 탐지하고, 보다 확실한 탐지를 위해 얼마나 올바르게 피부 병변을 탐지했는지 그 확률 값을 알 필요성이 있다고 판단하였다. 따라서 YOLOv8 객체 탐지(Object detection) 딥러닝 모델을 사용하였다. 또한 AI Hub에서 제공하는 반려동물 피부 질환 데이터 세트 중, 고양이 유증상 데이터를 활용하여 YOLOv8 모델을 학습시켜 고양이 피부 증상 탐지에 대한 정확도를 높이고자 했다.

1. 서 론

반려동물 인구가 1500만 명을 넘어서며, 반려동물 산업의 성장의 기반으로 반려묘를 대상으로 서비스를 고안하게 됐다. 특히 반려묘의 피부 질환은 초기 증상이 미미해 조기 진단이 어렵고, 병이 악화되는 경우가 많다는 문제에 주목했다. 피부 전문 병원의 접근성 부족, 질병 정보의 부족으로 인해 진단이 지연되는 경우가 빈번하다는 점도 확인됐다. 이를 해결하고자 AI 허브의 반려동물 피부 질환 데이터로 반려묘의 피부 증상을 확인해 조기 증상 탐지로 질병의 악화를 방지하며, 의료비 절감의 가능성이 있는 서비스를 개발했다.

2. 관련 연구

2.1 반려동물 품종 분류 연구

「딥러닝 YOLOv5 객체 인식을 이용한 반려동물 품종 분류 방법에 대한 연구」 [1]에서는 반려동물 품종 인식 개선을 위해 YOLOv5 알고리즘을 활용한 객체 인식 방법을 제안한다. 옥스퍼드 데이터 세트에서 수집된 반려동물 이미지를 이용하여 모델의 학습에 사용하였다. IoU 임계값을 설정하여 모델의 성능을 측정하였으며, 실험 결과는 적은 데이터량과 IoU 임계값 조정으로 반려동물 품종 인식의 성능이 약 30% 향상됨을 보여주었다.

3. 훈련 방식 및 선택 모델

3.1 Deep Learning

딥러닝은 머신러닝의 한 분야로, 인간 두뇌의 복잡한 의사 결정 시뮬레이션을 위해 심층 신경망이라고 불리는 다층 신경망을 사용한다[2]. 또한 심층 신경망은 상호 연결된 여러 계층의 노드로 구성되며, 각 계층은 예측 또는 분류를 구체화 및 최적화를 위해 이전 계층을 기반으로 한다. 이런 요소는 함께 작동해 데이터 내의 개체를 정확히 인식, 분류, 설명한다[2]. 따라서 많은 특징들을 포함하는 피부 병변 이미지 학습을 위해 딥러닝을 채택했다.

3.2 YOLOv8

이미지 내에서의 피부 병변을 탐지하고, 보다 더 확실한 탐지를 위해 얼마나 올바르게 피부 병변을 탐지했는지 그 확률 값을 알 필요성이 있다고 판단하였다. 따라서 YOLOv8 객체 탐지(Object detection) 모델을 사용하였다.

우선적으로 YOLO 모델은 You Only Look Once의 약자로, 이미지를 한 번만 보고 어떤 물체가 어디에 존재하는지 예측할 수 있다는 것이 큰 특징이다. 이는 즉, 단일 합성곱 네트워크(Single convolutional network)가 여러 개의 경계 상자와 해당 상자에 대한 클래스 확률을 동시에 예측한다는 것이다. 따라서 YOLO 모델은 전체 이미지에 대해 학습하고 탐지 성능을 직접 최적화한다[3]. 이 중 YOLOv8은 이전 YOLO 버전의 발전된 기능을 바탕으로 새로운 기능과 최적화가 도입된 최신 버전 모델이다 [4]. YOLOv8에서의 출력 이미지는 탐지된 객체의 박스 윤곽선, 클래스 이름, 신뢰도 점수가 함께 표시된다.

4. 실험 및 결과

4.1 데이터 전처리

AI Hub의 반려동물 피부 질환 데이터 중 고양이 농포, 여드름(A4), 결절, 종괴(A6) 두 클래스를 사용해 총 9600장 유 증상 이미지를 학습(8000장)과 검증(1600장) 데이터로 8:2 비율로 나눴다. 이미지 크기는 640x640으로 통일하고, 좌표 값도 비율에 맞춰 재조정했다. bbox 라벨링 값을 YOLO 방식의 좌표 값으로 변환했으며, polygon 좌표만 있는 경우엔 bbox를 추출해 YOLO 형식으로 변환했다. 라벨링은 클래스 번호, 중심 좌표(x, y), 너비, 높이 순서로 이루어지며, 좌표는 0~1 사이로 정규화된다.

4.2 모델 튜닝

딥러닝 학습에는 YOLOv8 모델을 사용했으며, 보다 더 나은 결과를 위해 사전 학습된 모델을 사용했다. 모델 훈

런 시에는 파이토치(PyTorch) 프레임워크를 사용했다. 최적화 방법은 확률적 경사 하강법을 사용했고, 과대적합 방지를 위해 드롭아웃과 조기 종료법을 적용했다.

4.3 실험 과정 및 결과

1차 시도에선 전이 학습 모델을 사용하지 않고 시도했다. 훈련 데이터는 총 1000장, 검증 데이터는 총 200장을 사용했다. epochs는 20, learning rate는 0.01, image size는 640x640, batch는 16으로 설정했다.

(표 1) 1차 시도 결과

Class	mAP50
all	0.0353
A4	0.0522
A6	0.0185

현저히 낮은 결과가 나와 이후 시도에서는 전이 학습이 적용된 모델로 시도했다. 설정한 하이퍼 파라미터는 1차 시도와 같다. 적용 후 훈련한 모델에서는 1차 시도 때보다 정확도가 오른 것을 확인할 수 있었다.

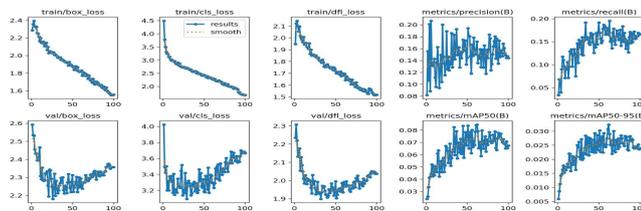
(표 2) 2차 시도 결과

Class	mAP50
all	0.0692
A4	0.0786
A6	0.0597

이전 실험에서 정확도가 너무 낮게 나와 데이터 세트를 추가 후, 하이퍼 파라미터를 조정했다. 훈련 데이터 2000장, 검증 데이터 400장을 사용하였고, epochs는 100, learning rate는 0.001, batch는 8로 조정했다.

(표 3) 3차 시도 결과

Class	mAP50
all	0.0848
A4	0.101
A6	0.0685



(그림 1) 3차 시도 결과 그래프

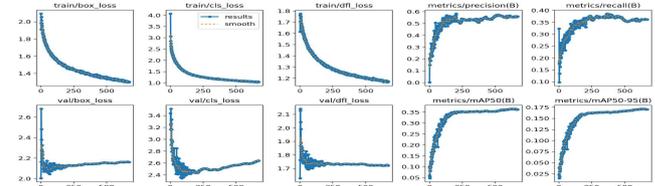
실험 결과, 이전 시도보다는 정확도가 근소하게 상승했으나 검증 데이터 세트에서 epochs이 약 40번째일 때부터 과대 적합이 일어나는 것을 확인할 수 있었다. 따라서 이후 실험에서는 드롭아웃 및 조기 종료법을 적용했다.

최종 실험에서는 훈련 데이터 8000장, 검증 데이터 1600장을 사용했다. epochs는 2000, learning rate는 0.01, batch(auto로 조정하였다.)는 113, image size는 640으로 두었다. epochs를 늘린 대신 patience 값(개선이 되지 않더라도 바로 학습을 종료하지 않고 지켜보는 epoch 수[4].)를 50으로 두었다. 또한 과대적합을 방지하기 위해 드롭아웃은 0.3으로 적용했다. 그 결과 총 670 epoch이 수행되었으며, mAP50 지수(IoU 임계값 0.50~0.95 범위에서 계산된 평균 정밀도의 평균이다 [4].)는 이전보다 비교적 크게 오른 것을 확인했다. 또한 훈련 데이터에서의 loss 값도 안정적으로 줄어드는 것을 볼 수 있었다. 그러나 검증 데이터에서의 box loss 값과

class loss 값이 일정 epoch 수부터는 증가하는 양상을 보아 모델이 검증 데이터에 과대 적합 되었음을 알 수 있다.

(표 4) 최종 결과

Class	mAP50
all	0.364
A4	0.169
A6	0.559



(그림 2) 최종 결과 그래프

최종 모델로 테스트 결과, 판별된 피부 병변의 클래스 이름과 판별 확률이 box 형태로 표시됨을 확인했다. 해당 예측 사진에서는 A4(농포와 여드름) 병변이 측정됐다.



(그림 3) 농포와 여드름 병변 측정 사진

최종 시도에서 전체 mAP50 지수는 0.364로 측정되었으나, 실제 상용화하기에는 어려운 수치다. 따라서 향후 연구에서는 더 다양한 병변 데이터 추가와 batch size, learning rate를 특정 비율로 조정하는 등, 하이퍼 파라미터 최적화를 통해 개선할 것이다. 이를 통해 추후 가정에서도 반려동물의 병변을 빠르게 식별, 대처할 수 있을 것이라 판단된다.

참고 문헌

[1] 박태주, 송특섭, "딥러닝 YOLOv5 객체 인식을 이용한 반려동물 품종 분류 방법에 대한 연구", 한국정보통신학회논문지, 제27권, 제4호, pp. 565-570, 2023년 4월
 [2] 딥러닝이란 무엇인가요?(2024). <https://www.ibm.com/kr-ko/topics/deep-learning> (accessed Aug., 22, 2024).
 [3] You Only Look Once: Unified, Real-Time Object Detection(2016). <https://arxiv.org/abs/1506.02640> (accessed Aug., 9, 2024).
 [4] Ultralytics YOLO Docs(2023). <https://docs.ultralytics.com> (accessed Aug., 12, 2024).

※ 본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음 (2024)

유학생을 위한 문화 교류 매칭 및 커뮤니티 서비스 구현

윤부영, 강예림, 김수연, 황준
서울여자대학교 소프트웨어융합학과, 불어불문학과, 생명환경공학전공
e-mail : yyd95060@gmail.com

Implementation of cultural exchange matching and community services for international students

Buyoung Yoon, Yerim Kang, Suyeon Kim, Jun Hwang
Department of Software Convergence, Seoul Women's University

요 약

타지에서 온 유학생들 사이에서 기존의 거리 기반 소셜 애플리케이션들을 통한 새로운 만남은 안전성과 신뢰성의 문제가 지속적으로 제기되어 왔다. 이에 본 연구는 학생증 인증과 키워드 검색 필터링을 통해 자신의 목적에 부합하는 친구들과 안전하게 교류하며 소통할 수 있도록 해 이러한 문제들을 극복하고자 하였다. 회원가입 과정에서 학생증 인증을 통해 안전하게 서비스를 이용할 수 있도록 하였고, 매칭 시스템과 해시태그 및 조건 필터링 기능으로 사용자들에게 맞춤 서비스를 제공함으로써 목적에 맞는 교류와 정보 공유를 가능케 하였다.

1. 서 론

인터넷의 발달과 국제화로 인해 한국으로 유학하러 오는 국제 학생들이 늘어나고 있음에도 불구하고, 언어 문제와 더불어 만남의 기회가 부족하여 유학생으로서 친구를 사귀는 데 큰 어려움을 겪고 있는 현실이다. 외국의 상황도 크게 다르지 않다. 한인 커뮤니티 형성의 어려움과 현지인과의 교류 부족 문제는 꾸준히 제기되고 있다. 기존의 다양한 사람들의 만남을 지원하는 거리 기반 소셜 애플리케이션(PNA: People Nearby Application)을 통하여 새로운 만남을 기대해 볼 수 있으나 이것에 대한 신뢰는 아직 부족한 상황이다. 또한 현재 널리 사용되고 있는 데이트 애플리케이션 'Tinder'는[1] 지속적인 관계가 아닌 캐주얼하고 성적인 만남을 위한 애플리케이션으로 기능한다는 인식이 많다.[2] 이를 극복하기 위해 만들어진 독일의 서비스는[3] 회원들의 가입 신청서를 개별적으로 확인하고 있지만, 별도의 신분 확인 절차가 없다는 문제점이 존재한다. 이에 본 연구에서는 매칭 애플리케이션들이 가진 안전성과 신뢰성 문제를 극복하고 유학생들에게 안전한 만남의 장을 제공하기 위해 본 애플리케이션을 개발하게 되었다.

본 연구에서는 국제 학생뿐만 아니라, 외국어를 배우고 외국인 친구를 사귀고 싶어 하는 한국 학생의 요구도 수용하여 대학생들이 널리 사용할 수 있는 시스템을 제안한다. 해당 서비스는 이용자의 세부 정보 입력과 학생증 인증을 통하여 신원을 보장할 수 있도록 하였으며, 조건 검색 필터링을 통해 이용자의 목적에 부합하는 친구들을 사귄 수 있는 서비스를 제공한다.

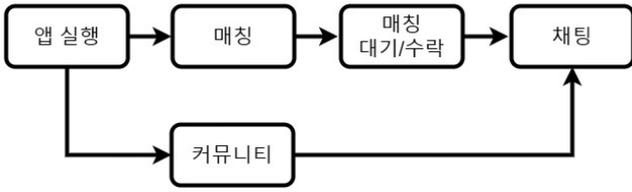
2. 서비스 설계 및 구현

2.1 서비스 개요

본 연구에서 개발한 서비스는 iOS와 Android 플랫폼에서 모두 사용이 가능하도록 Dart 언어를 활용하여 개발되었다. Firebase 서버를 이용하여 회원 정보 저장이 가능하도록 하였다. 이를 통해 사용자 데이터를 안정적으로 관리할 수 있다. 나아가 사용자의 검색 의도와 목적에 따라 원하는 상대방을 더욱 잘 찾을 수 있도록 안정 매칭 알고리즘인 Gale-Shapley 알고리즘을 선정하였다.

전통적인 Gale-Shapley 알고리즘은 안정적인 매칭을 제공하는 장점이 있으나, 참가자의 개별 선호가 충분히 반영되지 않거나 특정 상황에서 매칭 만족도가 낮아질 수 있다는 한계가 존재한다. 따라서 본 연구에서는 학생들의 세부 정보를 활용한 맞춤형 필터링 기능을 통해 이러한 단점을 보완하고, 안정적인 매칭의 장점을 극대화하여 만족스러운 교류가 이루어지도록 설계하였다.

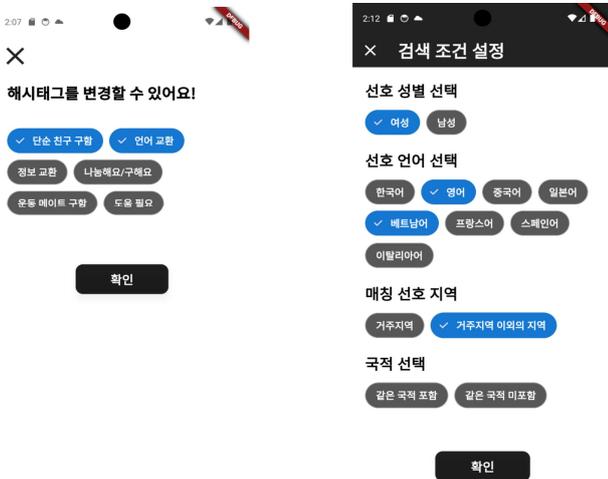
본 애플리케이션에서 지원하는 서비스 절차는 그림 1과 같이 크게 2가지 경로로 나눌 수 있다. 첫 번째 서비스는 사용자가 원하는 조건으로 검색하여 원하는 사용자와 이루어지는 매칭 서비스이다. 상대방이 매칭을 수락하고 나면 1:1 채팅이 가능하다. 두 번째 서비스는 커뮤니티를 이용하는 다수의 사용자와 이루어지는 교류 서비스이다. 1:1 채팅뿐만 아니라 커뮤니티를 이용하는 이용자 간의 채팅도 가능해 다양한 문화적 교류가 가능하다.



(그림 1) 서비스 흐름도

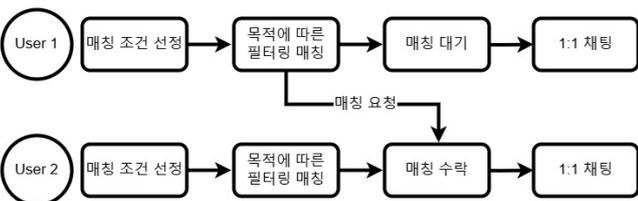
2.2 서비스 주요 기능

서비스의 주요 기능 중 하나는 회원가입 시 학생증 사진을 첨부하여 학생임을 인증하는 것이다. 첨부된 학생증 사진은 운영자가 확인한 후, 사용자에게 매칭 및 1:1 채팅 기능을 사용할 수 있는 권한을 부여한다. 이를 통해 서비스의 신뢰성을 확보하고, 안전하게 서비스를 이용할 수 있도록 하였다.



(그림 2) 매칭 검색 필터링을 위한 조건 설정 화면

학생임을 인증받은 다음, 사용자는 그림 2에서의 프로필을 작성할 때 해시태그를 설정할 수 있다. 해시태그를 선정하는 데 어려움을 느끼는 사용자들을 고려하여 기본적으로 제공되는 해시태그를 사용할 수 있게 하였고, 제공된 해시태그가 자신을 제대로 설명하지 못한다고 느끼는 사용자의 경우에는 직접 해시태그를 작성할 수 있도록 하여 자유도를 높였다. 뿐만 아니라 성별, 거주 지역 등 다양한 기준으로 친구를 쉽게 찾을 수 있도록 도와주는 검색 기능도 서비스의 주요 기능 중 하나이다. 검색 기능을 통해 사용자는 자신의 의도와 목적에 맞는 친구를 더욱 편리하게 찾을 수 있다.



(그림 3) 매칭 흐름도

또한, 그림 3과 같이 사용자는 매칭 요청 기능을 통해 다른 사용자에게 매칭을 요청할 수 있으며, 상대방이 매칭을 수락해야 1:1 채팅이 시작될 수 있도록 하였다.

마지막으로는 커뮤니티 게시판을 활용하여 일대일 소통뿐 아니라 다수가 함께 소통할 수 있도록 지원하는 1:n 소통 기능이다. 이러한 기능을 통해 사용자들은 다양한 의견을 공유할 수 있으며, 많은 사용자들과 다양한 의견을 공유할 수 있다.

이와 같은 기능들을 통해 본 서비스는 사용자에게 신뢰성을 높이고, 상황에 맞는 인연들을 편리하게 찾을 수 있도록 도움을 주며, 한 가지 방식이 아닌 다양한 방식으로 소통할 수 있는 환경을 제공한다.

3. 결론

본 연구에서는 학생증 인증과 더불어 매칭 요청 거절 기능을 통해 학생들 사이에서의 안전한 만남의 장을 제공하는 매칭 애플리케이션을 고안하였다. 해당 애플리케이션은 해시태그 및 키워드 필터링 기능으로 사용자들을 필터링하는 맞춤형 서비스를 제공함으로써 사용자의 목적에 맞는 교류와 정보 공유를 가능케 한다.

이를 통해 국제 학생들과 한국 학생들 간의 활발한 교류를 통해 진정한 글로벌 사회를 경험할 수 있는 풍부한 학교생활을 기대할 수 있다. 이러한 상호작용은 다양한 문화적 배경을 가진 학생들이 서로 배우고 성장할 기회를 제공하여, 세계 시민으로서의 역량을 함양하는 데 기여할 것이다.

다만, 학생증 인증의 경우에는 관리자가 확인 후 승인하는 방식의 시스템이기 때문에 신원 도용, 학생증 위조 등의 문제가 생길 가능성이 있으므로 이러한 점을 보완하기 위한 추가적인 연구가 필요하다.

ACKNOWLEDGMENTS

“본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)“

참고 문헌

[1] Tinder LLC, Tinder(버전 15.16.0) [모바일 앱] Google Play Store
 [2] 김효림(Hyolim Kim), 오서림(Seolim Oh), 이가은(Gaeun Lee), 정수민(Sumin Jung), 정규하(Kyuha Jung), 임하진(Hajin Lim), “데이팅 앱의 사용 동기의 변화와 그에 따른 사용 양상의 변화 - Tinder 국내 사용자 경험을 중심으로”, 한국HCI학회 Vol.2023 No.2 [2023]
 [3] Tripod Technology GmbH, Tandem(버전 5.34.0) [모바일 앱] Google Play Store

콤팩 : 이주 외국인과 선생님을 위한 한국어 숙제 지원 애플리케이션

이윤지*, 석민솔*, 신지우**, 박시영*, 엄성용*
 서울여자대학교 소프트웨어융합학과*, 서울여자대학교 디지털미디어학과**
 e-mail : yoonji01ee@gmail.com, msseock@gmail.com, sjw01827@gmail.com,
 psy915@swu.ac.kr, osy@swu.ac.kr

KoApp : Korean homework support application for migrant foreigners and teachers

Yun Ji Lee*, Mihn Sol Seock*, Ji Woo Shin**, Si Yeong Park*, Seong Yong Ohm*
 Department of Software Convergence*, Department of Digital Media Design and Applications**, Seoul Women's University, Korea

요 약

이주 외국인을 위한 AI 기반 한국어 숙제 지원 애플리케이션을 기획하여 기존 한국어 수업의 부족한 부분을 보완하고, 한국어 학습의 흥미 유발 및 효율성을 높이는 것을 목표로 한다. 앱은 힌트, 다국어, 자동 채점, AI 피드백 기능 등을 제공하여 학습효과와 교수자의 편의 증대를 기대할 수 있다. 향후 다양한 숙제 템플릿을 추가하여 사용자 만족도를 높이고자 한다.

1. 서 론

현대 사회의 글로벌화로 인해 한국에서도 이주 외국인의 수가 지속적으로 증가하면서, 언어 장벽 문제가 중요한 사회적 문제로 부각되고 있다. 2022년 기준 국내 등록 외국인[1]은 약 118만 명을 넘었으며, 이들이 한국 사회에 적응하고 일상생활에서 원활히 소통할 수 있도록 돕는 것이 중요한 과제가 되었다. 현재 서울의 다문화 가족 센터에서는 한국어 학습 프로그램을 운영하고 있지만[2], 통역 교사 배치, 숙제 출제 및 이해 어려움 등 차질을 겪고 있다. 이주 외국인에 대한 체계적인 교육의 시행이 어렵고 관리마저 소홀하여 교육적 한계 또한 존재한다[3]. 이에 본 연구는 이러한 실정을 바탕으로 이주 외국인 학습자를 위한 한국어 숙제 지원 애플리케이션을 개발하고자 한다. 특히 인공지능을 활용한 말하기 연습은 언어 발달과 자기주도 학습에 효과적이다[4]. 따라서, 인공지능 기반의 한국어 학습 도구를 적용하여 한국어 교수자의 부담을 줄이고, 학습자에게 더 나은 학습 경험을 제공할 것을 목표로 한다.

2. 관련 연구

이주 외국인을 위한 한국어 학습 애플리케이션은 존재하지 않지만, 이주 외국인을 대상으로 한 숙제 애플리케이션은 찾아보기 어렵다. 그러나 본 연구에서 기획한 '콤팩' 애플리케이션은 이주 외국인을 고려한 힌트 기능과 다국어 기능이 있고, 교수자를 고려한 자동 채점 기능, 쓰기 숙제에 대한 AI 피드백 기능과 숙제 템플릿 선택 기능이 있어 사용자의 편의성을 증대시킨다.

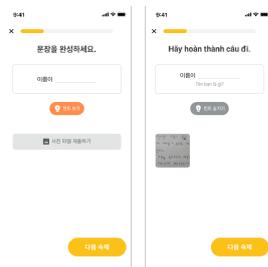
3. 주요 기능 및 화면 구성

3.1 숙제 힌트 기능

이주 외국인은 한국어에 서툴러, 선생님이 내준 숙제를 잘

이해하지 못하는 경우가 많다. 이때 사용할 수 있는 기능이 숙제 힌트 기능이다. 그림 1과 같이, 힌트 보기 버튼을 클릭하면, 문제의 지문을 학습자의 디바이스 기본 언어로 번역하여 보여준다. 이때 지원되는 번역 언어는 한국어 거주하는 이주 외국인 통계[2]에 따라 가장 많이 사용되는 베트남어 또는 중국어이다. 사용자가 디바이스 기본 언어로 베트남어 또는 중국어를 설정해 두었다면, 문제의 질문 또한 해당 언어로 번역하여 보여준다. 학습자는 힌트 기능과 다국어 기능을 통해 문제 이해도를 높여 한국어 숙제를 푸는 데 도움을 받을 수 있다.

특히 쓰기 숙제의 경우, 학습자는 쓰기 숙제를 풀 때 하나의 사진 파일만 올릴 수 있다. 학습자가 제출한 사진 파일은 교수자가 숙제를 검사할 때 확인할 수 있다.



(그림 1) 학습자의 쓰기 숙제 화면

3.2 한국어 말하기 연습 기능

이주 외국인은 한국에서 생계를 이어가기 위해 일상에서 한국어를 사용해야 한다. 2022년 기준, 국내 등록 외국인의 성비는 큰 차이가 없으므로[5], 성별에 관계없이 일상생활 속 자주 사용되는 6가지 주제인 '가족', '학교', '날씨와 계절', '여행', '쇼핑', '미디어와 콘텐츠'를 말하기 연습 주제로 선정했다.

학습자는 선택한 주제와 상황에 맞게 인공지능과 말하기 연습을 할 수 있다. 힌트 기능을 통해, 인공지능의 질문에 대한

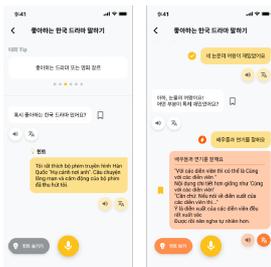
적절한 답변을 추천받을 수 있다. 사용자가 문법적으로 올바른 답변을 했을 경우 말풍선 왼쪽에 체크 마크가, 그렇지 않은 경우 번개 마크가 표시된다. 번개 마크가 표시되면 수정된 문장과 이유를 제공한다.

이 과정에서 그림 2와 같이 Google의 생성형 AI인 Gemini 1.5 Flash로 프롬프트 엔지니어링을 활용하여 주제별 대화 모델, 답변 힌트 제공 모델, 문법 검사 및 피드백 모델을 제작했다. 모든 프롬프트에는 이주 외국인도 쉽게 이해할 수 있는 문장으로 답변하도록 지시하여, 한국어 학습의 어려움을 최소화하고자 했다.



(그림 2) 한국 드라마 말하기 주제 프롬프트

아래 그림 3은 좋아하는 한국 드라마 말하기를 주제로 한국어 말하기 연습을 하는 과정을 담고 있다. 첫 번째 이미지에서 사용자가 '힌트 보기' 버튼을 누르면, 힌트 모델이 좋아하는 한국 드라마 질문에 대한 자연스러운 대답을 이용자의 모국어로 보여주며, 답변을 유도한다. 이어서 사용자가 '내 눈물의 여왕이 재밌었어요'라고 말하면 문법 검사 모델이 문장을 적절하다고 판단해 말풍선 왼쪽에 체크 마크를 보여주고, 대화 모델은 작품에서 재밌던 부분을 묻는 말로 이어간다. 만약 사용자가 '배우들과 연기를 잘해요.'처럼 대화의 흐름에 적절하지 않은 문장을 말하면, 문법 검사 모델이 올바른 문장과 그 이유를 이주 외국의 눈높이에 맞게 제시한다. 이와 같은 과정의 말하기 연습을 진행하면서 외국인의 한국어 학습을 돕는다.



(그림 3) 학습자의 말하기 연습 화면

3.3 속제 제공 기능

교수자는 학습자에게 듣기, 말하기, 읽기, 쓰기, 놀이형 속제 총 5가지 유형의 속제를 낼 수 있다. 듣기 속제는 스피커로 한국어 문장을 듣고 선택지에서 답을 선택하는 방식이다. 말하기 속제에서는 주제나 지문을 보고 학습자가 녹음한 오디오 파일을 제출한다. 읽기 속제는 지문을 읽고 녹음한 오디오 파일을 제출하는 형태이다. 쓰기 속제는 지문을 보고 한국어 문장을 완성하는 것이다. 놀이형 속제는 숨은 낱말 찾기, 다섯 고개 등이 있다. 다양한 속제 템플릿은 교수자가 간편하게 속제를 제공할 수 있게 하고, 학습자는 더 흥미롭게 학습할 수 있도록 돕는다.

3.4 속제 검사 기능

교수자는 학생이 제출한 속제에 대하여 피드백을 남길 수 있다. 쓰기 속제의 경우 사용자가 속제로 제출한 사진이 있다면, 사진 내의 문장을 OCR 기능으로 텍스트화하여 텍스트화한 문장을 인공지능으로 맞춤법 검사 및 띄어쓰기 교정을 할 수 있다. 교수자는 쓰기 속제 사진을 보고 채점하는 시간이 줄어서 편의성이 증대된다.



(그림 4) 교수자의 속제 검사 화면

4. 결론 및 향후 연구 방향

본 애플리케이션 '콧'은 이주 외국인을 위한 한국어 교육 앱으로, 속제 기능과 말하기 연습 기능을 통해 한국어 능력 향상에 다방면으로 이바지하고자 기획되었다. 현재 이주 외국인을 대상으로 한 한국어 교육 현장에서 활용할 수 있는 속제 애플리케이션이 부재하다는 점에서 타 서비스들과 차별성을 지닌다. 번역 힌트를 통해 언어적 차이를 해소하며, 말하기 연습 기능을 통해 시간과 장소에 구애받지 않고 말하기 연습을 할 수 있다는 점에서 사용자들에게 편의성을 제공한다. 향후 현재 개발된 기능 외에 더 다양한 속제 템플릿을 제공하여 교수자에게는 폭넓은 선택지를, 학습자에게는 다양한 학습 방식을 제공한다면, 보다 많은 사용자가 만족하며 활용할 수 있는 애플리케이션으로 자리매김할 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 서울여자대학교 SW중심대학추진사업단의 지원의 연구결과로 수행되었음(2024)

참고문헌

- [1] 법무부. “출입국자 및 체류외국인 통계”, 지표누리, 2023. <https://www.index.go.kr/unity/potal/indicator/IndexInfo.do?cdNo=210&clasCd=10&idxCd=F0013&upCd=1>, (accessed Sep., 11, 2024).
- [2] 한국건강가정진흥원. “2022 가족지원사업 연간 결과 보고서”, 2022.
- [3] 이유인. “한국 정부의 이주노동자 대상 한국어교육 정책 연구”, 국내박사학위논문 상명대학교 일반대학원, 2018.
- [4] 장지혜, 김정렬. “AI펍톡을 활용한 초등영어 말하기 수업의 효과”, 『초등교과교육연구』 제37호, 2022. pp.37-55.
- [5] 법무부. “성 및 연령별 등록외국인 수”, 지표누리, 2024. <https://www.index.go.kr/unity/potal/indicator/IndexInfo.do?cdNo=210&idxCd=F0014>.

디지털트윈 기반 파이프랙 강구조물의 동특성 모니터링 시스템 구축

송민호, 이서진, 양재영, 김현우, 박재용, 최재혁*
조선대학교 건축공학과, (주)스마트오션

e-mail : smh9337@naver.com, sangsuri1004@naver.com, jy_yang1@naver.com, jerry7237@naver.com, parkqwqwee99@smartocean.co.kr, jh_choi@chosun.ac.kr

Establishment of a dynamic characteristic monitoring system for pipe rack based on digital twins

Minho Song, Seojin Lee, Jaeyoung Yang, Hyunwoo Kim, Jaeyoung Park, Jaehyouk Choi*
Department of Architectural Engineering, Chosun University

요약

본 논문에서는 디지털트윈 레벨 3(Simulation)을 구현하기 위해 현장 구조물에 가속도 계측기를 설치하고 측정된 데이터를 원거리통신네트워크를 활용하여 가상공간 내에서 기존 구조체의 동적특성, 특히 구조물의 고유진동수 및 감쇄특성을 파악하기 위한 현장 모니터링 시스템 구축에 대하여 검토하였다.

1. 서론

주로 국가산업단지 내에 설치된 파이프랙(Pipe-rack) 구조물은 파이프라인(Pipeline)을 지지하기 위한 강구조물로 형성되어 있다. 플랜트 시설물과 같은 대규모 산업시설의 파이프라인은 생상품목의 확충 등 그 필요성에 따라 잦은 증축과 보강이 이루어지고 있다. 기존의 노후화된 파이프랙의 경우 사용연한이 오래됨에 따라 콘크리트 기초부의 열화, 지반침하 및 강재의 부식 등의 결함으로 기존의 구조안정성에 대한 검토가 요구되고 있다.

최근, 디지털트윈 기반의 시설물 안전점검 및 유지관리에 관한 연구가 활발히 진행되고 있다. 디지털트윈은 현장의 상황이 가상공간에서도 그대로 미러링(Mirroring) 및 모니터링(Monitoring)이 가능하므로 유지관리 효율을 극대화할 수 있는 이점이 있다.

서 기존 구조체의 동적특성, 특히 구조물의 고유진동수 및 감쇄특성을 파악하기 위한 현장 모니터링 시스템 구축에 대하여 검토하였다.



(그림 2) 파이프랙 결함 현황

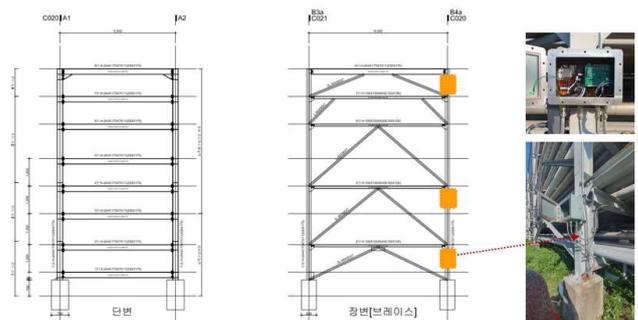


(그림 1) 파이프라인 및 파이프랙 구조물

본 논문에서는 디지털트윈 레벨 3(Simulation)을 구현하기 위해 현장 구조물에 가속도 계측기를 설치하고 측정된 데이터를 원거리통신네트워크를 활용하여 가상공간 내

2. 대상 구조물 및 가속도 센서의 설치 위치

파이프랙 구조물의 고유진동수를 파악하기 위하여 (그림 2)에서 나타낸 바와 같이 7층 구조물의 하층부(1층), 중층부(3층) 및 상층부(7층)에 가속도계를 설치하였다.



(그림 3) 파이프랙 구조물의 현장 가속도계 설치위치

2.1 가속도계의 제원

계측에 사용된 3개의 무선가속도 센서(HWT605-485)는 샘플링 주파수 100Hz, 측정범위 0.1cm/s ~ 1,500cm/s의 가속도를 측정할 수 있도록 설정하였다.



(그림 4) 가속도 센서 및 무선통신모듈

3. 현장 강제진동 실험

3.1 현장 인력가진에 의한 강제진동실험

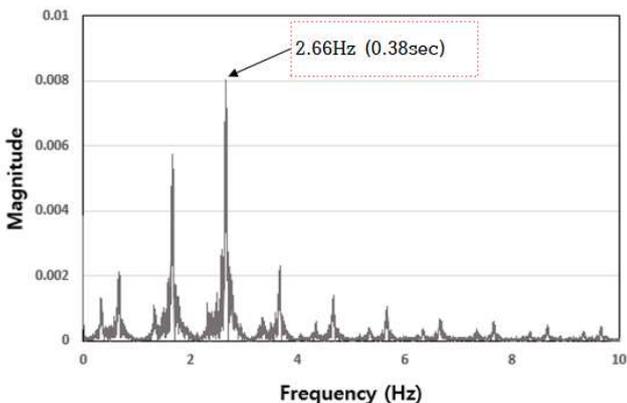
파이프랙 구조물의 강제진동실험을 위하여 안전한 접근이 가능한 6층 슬래브 하부에서 약 180Kg의 하중이동에 따른 수평가진을 실시하였다. 인력가진에 의한 공진현상을 유발하기 위해 예비 구조해석에서 얻어진 예상 공진주파수(1차모드, 약 1.67Hz)가 가능한 재현될 수 있도록 3회 가진하였다.

3.2 강제진동 실험결과

현장 강제진동 실험결과에는 (표 1)에서 보는 바와 같이 평균 2.6Hz로 나타나 예비 구조해석에서 도출된 1.67Hz보다 다소 높은 것으로 나타났다.

(표 1) 강제진동 실험 결과의 평균

구분		진동수 f (Hz)	주기 1/f (sec)	평균
순번	위치			
TEST 1	상층부	2.66	0.38	2.66Hz (0.38sec)
	중층부	2.66	0.38	
	하층부	2.66	0.38	
TEST 2	상층부	2.67	0.37	2.67Hz (0.37sec)
	중층부	2.67	0.37	
	하층부	2.67	0.37	
TEST 3	상층부	2.49	0.40	2.49Hz (0.40sec)
	중층부	2.49	0.40	
	하층부	2.49	0.40	

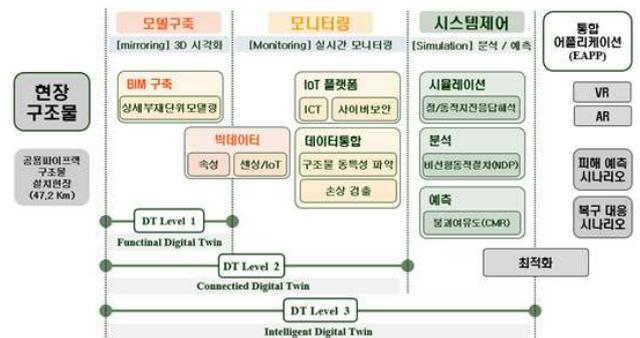


(그림 5) 강제진동실험 결과의 FFT해석 결과

이는 파이프랙의 특성상 파이프 라인이 상호 연결되어 있어 예비 구조해석에서는 반영이 어려운 파이프 라인 관의 연결로 인한 프레임 강성의 증가인 것으로 판단된다.

3.3 디지털트윈 가상영역에서 데이터 통합

실시간 모니터링 시스템 구축으로 현장에서 측정되는 가속도 데이터가 무선네트워크통신을 통하여 디지털트윈 가상공간으로 가속도데이터가 실시간으로 전송이 되고 있다. 또한 입력된 데이터를 기반으로 다양한 형태의 시뮬레이션(FFT 변환 등)이 가능함을 알 수 있고 (그림 6)의 시스템 제어의 DT Level 3의 시뮬레이션 및 분석에 해당하는 영역까지 구현이 가능하였다.



(그림 6) 디지털트윈 레벨 3: 시뮬레이션

4. 결론

구조물의 동특성을 파악하기 위해 현장에서 가속도계측 시스템을 구성하여 무선네트워크통신을 통하여 디지털트윈의 가상공간에 실시간으로 데이터를 전송하는 시스템을 구축하였다.

1. 인력가진 강제진동실험을 실시하여 구조물의 고유진동수를 측정된 결과, 단변방향에서 2.6Hz의 진동수를 확인하였다.
2. 이는 건축물 내진설계기준(KDS 41.17.00)에서 건축물의 유효주기를 산정하기 위하여 제시하고 있는 ' $T_a = C_t h_n^x$ ($C_t = 0.0724$, $x=0.8$: 철골모멘트 골조, h_n : 건축물의 밑면으로부터 최상층까지의 전체높이 (m))' 에 의한 2.17Hz(0.46sec)에 근접하고 예비해석결과값 1.67Hz보다는 다소 높은 값을 나타내고 있다.

감사의 글

이 연구는 2021년도 산업통산자원부 및 산업기술평가기관리원(KEIT) 연구비 지원에 의한 연구임(20017750)
이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2020R1A2C1009297)

참고문헌

- [1] 양재영, 권진영, 최재혁. (2024). 파이프랙 구조물의 상시 모니터링 기반 재난관리 플랫폼 구축 전략. (사)한국강구조학회 학술대회 논문집.
- [2] 시설물 영상취득 방안 체계화 및 손상분석 자동화 방안 연구 2023. 1. 국토안전관리원

디지털트윈 레벨3 구축을 위한 s-BIM 모델 생성과 실시간 모니터링

권진영, 허유정, 장현석, 김용주, 양재영, 최재혁*
 조선대학교 건축공학과

ds3403@naver.com huhyoul23456@gmail.com hs.jang@yunwoo.co.kr
 yjkim@corners.co.kr jy_yang@naver.com jh_choi@chosun.ac.kr

s-BIM modeling and real-time monitoring for building digital twin level 3

Kwon Jin young, Huh You Jeong, Choi Jae Hyouk*
 Chosun University Architectural Engineering

요약

파이프랙(Pipe-rack)은 파이프라인(Pipeline)시설의 일부로 주로 플랜트 시설물과 같은 대규모 산업 시설의 파이프 배관을 지지하기 위한 강구조물로 형성되어 있다. 현재 국가산단 내에는 설치 기간이 오래되어 노후화된 시설물이 많고 인명사고와 같은 중대재해가 빈번하게 발생하는데 이를 효과적으로 관리하기 위해서는 파이프랙 구조물의 유지관리를 위한 디지털트윈 기반 플랫폼이 구축이 필요하다.

본 연구에서는 강구조물인 파이프랙을 대상으로 디지털트윈 기반 상시모니터링 유지관리 플랫폼 구축 전략 및 현황에 대해 서술하고자 한다.

1. 서론

건축물의 3차원 공간정보를 생성하기 위해 라이더(LiDAR)나 3D스캐너 등 레이저 펄스를 활용하여 물체와의 거리 및 표면을 계측하고 3D 포인트 클라우드를 3D 공간정보로 조합하는 방법을 많이 사용하고 있다. 이와 같은 방식은 건축물 전체 구성 요소(보, 기둥, 설비 부분 등)를 객체화된 3차원 데이터가 아닌 단순 형상과 건축물대장 정보와 같은 표현적인 속성만을 제공하고 있어 재난시의 응답 거동 특성 파악이 요구되는 구조안전성 검토 등에서 사용이 제한적이다.

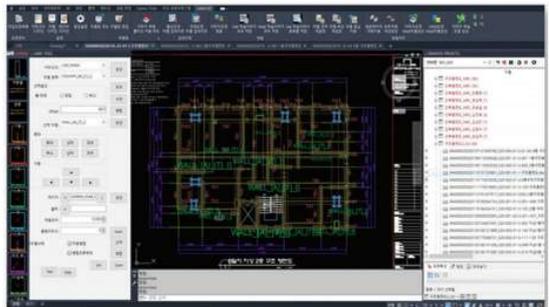
한편, BIM(Building Information Modelling)은 단순한 3D 모델링을 넘어 건축물의 각 구성 요소에 대한 속성정보를 점, 선, 면의 2D도면정보가 아닌 파라미터 기반 객체화된 데이터로 생성하여 관리하고 있다.

최근 공공부문 정보화 사업의 일환으로 전자정부 표준프레임워크 기반 건축물의 3차원 공간정보 구축에 관한 연구가 활발히 진행되고 있다. 기존 건축정보시스템(세움터 등)에서 단순 보유하고 있는 2D 전자도면에서 라벨링을 통한 인공지능 학습을 통해 형상정보(보, 기둥, 벽 등) 선별 인식/추출하여 3D s-BIM 디지털 모델이 개발되고 있다. 딥러닝 모델인 YOLO v3 기반으로 2D 도면의 구성요소를 현재 학습 수준에서 Detection의 경우 약 68%, 탐색된 이미지에 대한 Classification의 경우는 87%의 정확도로 3D 건축물 공간정보를 생성할 수 있음을 제시하였다.

본 논문에서는 디지털트윈 레벨3(Intelligent Digital Twin)를 구현하기 위하여 건축물을 구성하는 상세부재정보 및

하중 정보까지 모두 탑재된 Midas Gen의 모델링 파일(.mgt)을 BIM Telar 모델링 파일(.db)파일로 변환하여 구조체의 3D 공간정보를 자동으로 생성하였다. 또한, 이렇게 구축된 3D 공간정보를 기반으로 웹 대시보드 환경에서 IoT 플랫폼에서 취득한 동특성 데이터를 s-BIM 3D 공간정보에 인입된 상태에서 구조물의 동적특성을 시물레이션하는 프로세스를 소개하고자 한다.

AutoCAD기반 라벨링 자동화 도구



Mask R-CNN모델 탐지결과



(그림 1) AI기반 2D도면 BIM전환 시스템(2024)

2. 디지털트윈 레벨3 구현

디지털트윈(Digital Twin, DT)은 현장의 물리적 객체(건

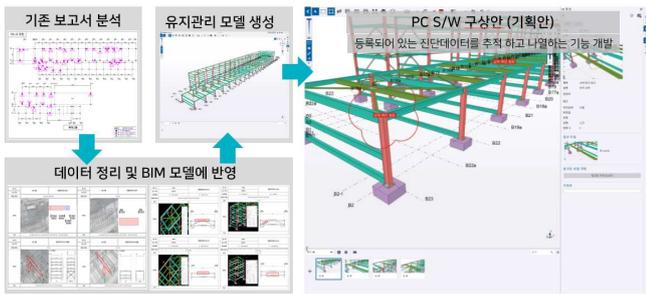
축물)가 컴퓨터 시스템의 가상 모델로 구현이 되고 이 두 부분이 실시간 데이터 전송으로 연결이 되어 있는 1단계 모사DT(Mirroring)와 2단계 관제DT(Monitoring)로 기본적으로 구성이 된다. 디지털트윈의 궁극적인 구현 목적은 가상공간 부분에서 시물레이션하는 모의DT(Modelling & Simulation) 3단계에서부터 실현된다고 할 수 있다. 설계자의 의도대로 자유롭게 시물레이션이 가능한 디지털트윈이 성공적으로 구축이 되면 이후 4단계 연합DT(Federaton)에서 다양한 형태의 실시간 시각화(가상현실/증강현실)의 구현 및 5단계 자율DT(Autonomous)에서 각종 시나리오별 의사결정에 강력한 도구로 활용이 가능하게 된다. 아래 (그림 2)에서 건축물 재난관리에 활용이 가능한 디지털트윈의 기술개발 프로세스를 나타내었다.

5단계	자율DT : Autonomous	재난시나리오별 최적대응방안 의사결정
4단계	연합DT : Federation	VR/AR을 활용한 통합어플리케이션 구현
3단계	모의DT : modelling & Simulation	정/동적해석을 통한 재난피해시물레이션 구축 머신러닝/시를 활용한 손상검출/분류
2단계	관제DT : Monitoring	센서/IoT 장치 실시간 데이터 취득
1단계	모사DT : Mirroring	s-BIM기반 도면/현장시공정보 가상공간 구축

(그림 2) 건축물 재난관리 디지털트윈 기술개발 프로세스

2.1 디지털트윈 레벨3 구현을 위한 BIM 모델 생성

BIM Model은 강구조물에 대해서 Tekla Structure 소프트웨어, LOD 350수준으로 기둥, 보, 브레이스, 접합부의 각 구성요소를 실측도면 기준으로 작성이 되었다. 기존의 Tekla Library에서 제공하지 않는 접합부 상세디테일은 사용자 component로 생성하고 등록된 프로파일을 통하여 동일 조건에서 반복 작업이 손쉽게 수행되도록 하였다. 본 연구에서 개발된 BIM 데이터 구축시스템은 (그림 3)에 나타내었다.

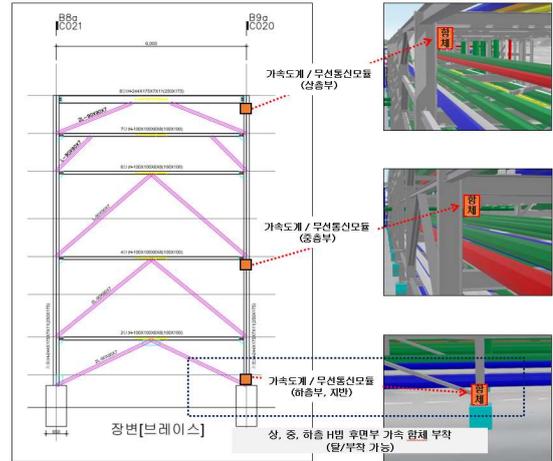


(그림 3) 현장 실측도면 기반 BIM 모델생성

2.2 디지털트윈 기반 실시간 모니터링(DT 2)

구조물의 동특성을 파악하기 위하여 실제 구조물에 가속도센서를 부착하여 웹소켓 네트워크 통신(WebSocket network communication)을 통하여 디지털트윈의 IoT플랫폼으로 데이터가 입력되도록 하였다. 이렇게 입력된 가속도 데이터는 가상공간의 BIM 모델링 데이터와 통합이 되어 구조물의 동특성을 파악하기 위한 FFT 시물레이션을 수행하게

된다.



(그림 4) 고속도계 실시간모니터링

3. 결론

디지털트윈 가상공간영역에 구성되어 있는 BIM 모델링과 가속도계의 계측데이터가 서로 통합되어 대쉬보드 상에 잘 표시되고 있는 것을 알 수 있다. 또한 가속도 데이터를 기반으로 FFT 시물레이션을 통해 구조물의 동특성(고유진동수 및 감쇄계수)의 변화 추이를 관찰함으로써 구조물의 건전성 유지 여부를 판단할 수 있다.



(그림 5) 대쉬보드 상에서 디지털트윈 레벨3의 구현

참고 문헌

[1] Zhao, Y., etc, (2020). A Deep Learning-Based Method to Detect Components from Scanned Structural Drawings for Reconstructing 3D Models. Applied Sciences, 10(6), 2066.
 [2] 양재영 etc, (2024), 파이프랙 구조물의 상시모니터링 기반 재난관리 플랫폼 구축 전략. 한국강구조학회 학술대회 발표집, 제주.
 [3] 김현우 etc, (2024), 디지털트윈을 통한 건축물의 재난 안전관리 프레임 워크 개발, 한국스마트미디어학회

감사의 글

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2020R1A2C1009297)

기기 에이전트 - 게이트웨이 배치 모델을 활용한 차등 프라이버시 적용 방안 연구

박시형
 중앙대학교 산업보안학과
 e-mail : a2514012@cau.ac.kr

A Study on the Application of Differential Privacy Using a Device Agent/Gateway-Based Deployment Model

Park Si hyung
 Dept. of Industrial Security, Chung-Ang University

요 약

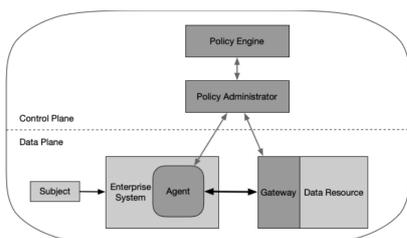
차등 프라이버시는 이론적으로 그 우수성이 입증되었음에도, 실제 구현의 어려움으로 인해 글로벌 IT기업을 제외하고는 많은 기관에서 활발히 사용되지 못하고 있다. 따라서 본 논문은 에이전트 - 게이트웨이 배치모델을 기반으로 한 차등프라이버시 프레임워크 구현 방안을 제안하여, 개인정보를 다루는 다양한 기관에서의 차등프라이버시의 채택과 활용을 장려하고자 한다.

1. 서 론

차등 프라이버시는 노이즈를 통해 원본데이터를 희석시킴으로써 데이터 프라이버시를 효과적으로 보장하는 개인 정보 강화 기술이다. 그러나 이러한 우수성에도 불구하고, 실질적인 차등 프라이버시 프레임워크를 구현하는 것이 쉽지 않기 때문에, 구글이나 애플과 같은 글로벌 IT기업을 제외한 많은 기관에서는 활발히 채택되지 못하고 있다[1][2]. 따라서 본 논문은 실질적이고 간단한 구현 방안으로 기기 에이전트 - 게이트웨이 배치 모델을 기반으로 하는 차등프라이버시 프레임워크를 제안하고, 높은 보안성을 입증하고자 한다. 이를 통해 데이터를 다루는 많은 기관에서 차등 프라이버시 기술이 활발히 채택되는데 기여할 수 있을 것으로 기대된다.

2. 기기 에이전트 - 게이트웨이 배치 모델

제로트러스트(Zero Trust)란, 내부망에 한번 접속을 성공하면 다양한 데이터에 별도의 인증 절차 없이 접근할 수 있는 경계 기반 모델의 한계를 보완하고자 대두된 보안프레임워크이다. 여기서 기기 에이전트 - 게이트웨이 배치 모델은 제로트러스트 프레임워크를 구현하는 대표적인 아키텍처 중 하나로, 내부망 내에 존재하는 데이터를 보호하는데 있어 효과적으로 사용될 수 있으며 자세한 구조는 아래와 같다[4].



(그림1) 기기 에이전트 - 게이트웨이 배치 모델 구조

해당 모델의 핵심 구성요소로는 에이전트, 게이트웨이, 그리고 정책관리자와 엔진(이하 ‘컨트롤러’) 으로 구분 할 수 있다. 먼저 에이전트란, 일종의 소프트웨어로서 접근주체의 기기에 설치되어, 사용자가 데이터를 조회하는 모든 접근이 해당 소프트웨어를 반드시 거치도록 한다. 게이트웨이는 데이터(또는 DB)의 흐름을 관리하는 소프트웨어로서, 컨트롤러의 명령에 따라 데이터의 접근을 허용하거나 거부하는 일종의 문지기 역할을 수행한다[5]. 컨트롤러는 에이전트로부터 접근주체의 신원정보와 접근하고자하는 데이터의 정보를 받아, 접근주체가 해당 데이터에 접근할 수 있는지의 신뢰여부를 최종적으로 결정하는 요소이다. 만약 컨트롤러가 접근주체의 데이터 접근을 허용하면, 에이전트와 게이트웨이에게 ‘데이터 채널’의 생성을 명령하여 접근주체가 데이터를 조회할 수 있게 한다. 정리하자면, 기기 에이전트 - 게이트웨이 배치 모델은 에이전트와 게이트웨이 그리고 컨트롤러 이 세 가지 구성요소가 서로 상호작용을 하여 궁극적으로는 높은 수준의 데이터보호를 달성할 수 있게 한다.

3. 차등 프라이버시

차등 프라이버시란, 2006년 C.Dwork에 의해 처음 제안된 개념으로, 쿼리의 응답에 수학적으로 정량화된 노이즈를 섞어, 원본 데이터의 특징을 유추할 수 없도록 하는 기술을 의미한다[6]. 정의에서부터 알 수 있듯이 차등프라이버시의 핵심요소는 노이즈로, 이는 프라이버시 예산이라는 변수에 의해 결정된다. 일반적으로 노이즈는 라플라스 분포 메커니즘에 의해 생성되는데, 프라이버시 예산의 값이 낮을수록 노이즈는 더 큰 범위에서 생성될 확률이 높고 반대로 프라이버시 예산의 값이 높을수록 노이즈는 0에 가까운 범위 내에서 생성된다는 특징이 있다. 그림2는 차등프라이버시를 통해 데이터프라이버시를 보장하는 과정을 보여준다.



(그림2) 차등프라이버시 적용 예시

이 경우 전송된 쿼리에 따라 원래 1이라는 결과가 나와야 하지만 +1이라는 노이즈가 추가됨으로써 최종적으로 2를 출력하였다. 이러한 노이즈는 연소득이 5,500만원 이하인 정보주체가 회원 탈퇴를 하더라도(데이터 셋에서 제거되더라도) 쿼리 결과의 차이가 크게 변화하지 않도록 하기 때문에 데이터 프라이버시를 보장할 수 있다.

4. 기기 에이전트-게이트웨이 배치 모델 기반 차등프라이버시 적용 시나리오

위와 같은 차등프라이버시가 기기 에이전트-게이트웨이 배치모델을 통해 구현되었을 경우의 데이터 접근 시나리오는 다음과 같다.

- ① 접근주체는 기기에 설치된 에이전트를 통해 특정 DB를 조회하고자 한다.
- ② 에이전트는 접근주체가 접근하고자하는 데이터와, 접근주체의 신원을 컨트롤러에게 전송한다.
- ③ 컨트롤러는 접근주체의 신원을 확인하고 그 신원이 해당 데이터에 접근하는데 적합한 경우 에이전트와 게이트웨이에 게 데이터채널을 생성하도록 명령한다.
- ④ 접근주체는 에이전트를 통해 쿼리를 입력하여 데이터 조회를 실시한다.
- ⑤ 접근주체가 전송한 쿼리 내용이 데이터채널을 통해 게이트웨이에게 전달되면, 게이트웨이는 DB에 접근하여 쿼리에 대한 응답을 생성한다. 이후 해당 응답에 적절한 프라이버시 예산을 사용하여 노이즈를 추가한다.
- ⑥ 게이트웨이는 노이즈가 추가된 최종적인 응답을 데이터 채널을 통해 에이전트에게 전송하며, 접근주체는 노이즈가 추가된 응답만을 확인할 수 있다.

이와 같은 절차는 컨트롤러를 통해 비 인가자의 데이터 접근을 근본적으로 제어하고, 인가받은 자가 쿼리를 통해 데이터에 접근하더라도 노이즈가 추가된 값만을 조회하게 함으로써 높은 수준의 보안성을 확보할 수 있을 것으로 기대된다.

5. 지능형 인공지능을 통한 추가 보안성 확보방안

5.1 게이트웨이의 프라이버시 예산 실시간 조절

게이트웨이는 데이터 앞에 배치되어 데이터에 대한 모든 흐름을 지속적으로 모니터링할 수 있다. 이때 만약 지능형 인공지능이 게이트웨이에 접목되어 특정 쿼리가 비정상적으로 반복되는 것과 같은 이상행위를 감지할 경우, 해당 쿼리가 데이터 추론 공격에 사용될 위험이 있음을 판단할 수 있다. 이에 따라 인공지능은 프라이버시 예산 값을 실시간으로 낮추어 노이즈가 크게 생성되도록 해 민

감한 정보가 추론될 위험을 실시간으로 줄일 수 있다.

5.2 에이전트의 쿼리 제어

에이전트는 게이트웨이와는 반대로 접근 주체의 기기에 직접 소프트웨어 형태로 설치되어 접근 주체가 자신을 통해서만 데이터에 접근할 수 있게 한다. 이 과정에서 마찬가지로 지능형 인공지능이 접목된 에이전트는 접근 주체가 전송하는 쿼리를 분석하고, 동일한 출력을 유도하는 쿼리가 비정상적으로 반복되는 것을 감지하면 쿼리의 입력 대기 시간을 생성하거나 차단하여 공격을 실시간으로 방지할 수 있을 것이다.

6. 결론

본 논문에서는 차등 프라이버시의 실질적 구현의 어려움을 해소하기 위해 기기 에이전트 - 게이트웨이 배치 모델을 활용한 프레임워크를 제안하였다. 해당 프레임워크는 쿼리를 전송할 때 마다 컨트롤러를 중심으로 지속적인 신원확인이 이루어지고, 쿼리에 대한 응답마저 차등프라이버시가 적용돼 노이즈가 추가되도록 하여 외부자는 물론 내부자의 위협까지 예방하는 훌륭한 수단으로 사용될 수 있다. 게다가 지능형 인공지능이 해당 과정에서 발생할 수 있는 이상행위까지 실시간으로 탐지, 대응하도록 한다면 이상적인 데이터 프라이버시를 보장할 수 있을 것으로 기대된다. 다만 본 연구는 단순 프레임워크만을 제안하였기 때문에 실제 환경에서 이러한 보안절차가 어떤 부정적인 영향을 끼칠 지를 파악하는 실험적 증거가 필요하다.

감사의 글

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

참고문헌

- [1] Graham Cormode, Somesh Jha, Tejas Kulkarni, Ninghui Li, Divesh Srivastava, Tianhao Wang. "Privacy at Scale: Local Differential Privacy in Practice." Proceedings of the ACM SIGMOD International Conference on Management of Data, vol. 18, no. 6, pp. 10-15, 2018.
- [2] OECD. "Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches." OECD Digital Economy Papers, no. 351, 2023.
- [3] 전보국, 최경택, 박규순. "보안은 불신으로부터, 제로트러스트 모델의 군 적용 필요성 제고." 한국방위산업진흥회, no. 551, pp. 116-129, 2021.
- [4] NIST. "Zero Trust Architecture." NIST Special Publication 800-207, 2020.
- [5] 한국인터넷진흥원, 한국제로트러스트포럼. "제로트러스트 가이드라인 1.0." 2023.
- [6] Cynthia Dwork. "Differential Privacy." Lecture Notes in Computer Science, vol. 4052, pp. 1-12, 2006.

대출가망 예측을 위한 랜덤포레스트와 차등 프라이버시의 융합: 안전한 데이터 활용 방안

이지용
 중앙대학교 산업보안학과
 e-mail : kevinoko@naver.com

Integrating Random Forest and Differential Privacy for Loan Default Prediction: A Secure Data Utilization Approach

Lee Ji Yong
 Dept. of Industrial Security, Chung-Ang University

1. 서론

대한민국은 국가 차원에서 금융 기관들이 금융 포용 정책을 적극적으로 추진하도록 장려하고 있다. 하지만 현실적으로 국민들이 금융 포용 정책을 수긍하지 못하고 있는 것이 현실이다. 본 논문은 대출가망고객에 대한 머신러닝 모델을 생성하여 직업 특성 상 신용점수가 낮아 대출 접근이 어려우나 대출이 필요한 사람들을 정확히 식별하여 적절한 대출 기회를 제공하고자 한다. 또한 안정성 측면에서 신용 위험이 높은 고객을 사전에 식별하여 금융기관의 부실 대출을 방지하여 대한민국의 경제 안정이 기여할 것을 기대한다. 개인은 적절한 맞춤형 대출 상품을 제안 받을 수 있어 과도한 대출로 인한 채무 불이행에 대한 위험을 방지할 수 있다. 하지만 대출 가능성이 있는 고객에 대한 데이터에는 개인의 주소와 직업과 같은 준식별자 정보, 전화번호와 같은 식별자 정보, 신용점수 및 연체 기록 등 민감한 개인정보가 포함되어 있다. 이러한 정보가 유출될 경우 개인뿐만 아니라, 국가와 사회에 막대한 비용을 초래할 수 있다. 따라서 본 논문은 차등 프라이버시 기술을 적용하여 데이터의 유용성을 유지한 채 개인정보가 식별되지 않도록 보호할 것을 기대한다.

2. 관련 연구

2.1 데이터 전처리

본 연구에서 데이터의 안전성을 보장하기 위해 대출가망고객 데이터에 대한 차등 프라이버시를 적용한다. 차등 프라이버시를 적용하기 이전에 데이터에 대한 전처리 과정을 거쳐야 한다. 해당 논문에서 데이터 전처리는 개인정보에 대한 가명처리를 의미한다. 나이 데이터는 그 자체로는 직접적인 식별자 정보가 아니지만, 충분히 세밀한 상태로 제공될 경우 준식별자로 작용하여 개인 식별의 위험성을 증가시킬 수 있다. 특히, 특정 나이 범위에 속하는 인구가 적은 경우, 해당 데이터와 다른 준식별자 정보(예: 성별, 거주지 등)를 결합하여 개인을 식별할 가능성이 커진다. 이러한 문제를 완화하기 위해 나이 데이터를 적절히 라운딩(rounding)하는 과정이 필

요하다. 거주지 데이터도 중요한 준식별자 중 하나로 세부적인 주소까지 제공될 경우 개인 식별의 위험성을 크게 증가시킨다. 이를 방지하기 위해 거주지 데이터를 시와 구 수준으로만 남기고 세부적인 주소 정보를 제거하여 개인정보 보호를 실현해야 한다. 데이터를 가명처리 한 후 라운딩한 나이 데이터와 성별 데이터를 기준으로 동질집합의 크기를 계산하고 최솟값을 k -익명성의 k 값으로 한다.

2.2 위험평가 기반 식별성이 높은 데이터 대상 비식별처리

k -익명성을 기반으로 한 위험성 평가를 기준으로 데이터를 정규화하여 정규화된 값에 따른 위험도 등급을 나눈다. 정규화 값은 해당 동질집합의 $(k - k_{\min}) / (k_{\max} - k_{\min})$ 을 통해 계산한다. k 는 해당 동질집합의 크기, k_{\max} 은 전체 데이터에서 가장 큰 동질집합의 크기, k_{\min} 은 전체 데이터에서 가장 크기가 작은 동질집합의 크기를 의미한다. k -익명성에 기반한 위험성 평가를 적용하여, 데이터를 정규화한 후 해당 정규화된 값에 따라 위험도 등급을 나누는 방식을 사용한다. 정규화 값은 다음과 같은 수식을 통해 계산된다.

$$\text{normalized value} = \frac{k - k_{\min}}{k_{\max} - k_{\min}}$$

여기서 k 는 해당 동질집합의 크기, k_{\max} 는 전체 데이터에서 가장 큰 동질집합의 크기, k_{\min} 은 가장 작은 동질집합의 크기를 나타낸다. 이를 통해 각 동질집합의 상대적 위험성을 정량화하고, 정규화된 값에 따라 위험도를 등급화하여 데이터를 보다 안전하게 처리할 수 있다. 위험도가 클수록 엡실론 값을 낮게 설정하여 차등 프라이버시를 실현한다. 표 1은 정규화된 값(normalized value)에 따른 위험도 측정과 엡실론 값 설정을 보여준다.

(표 1) 정규화한 값에 따른 위험도 측정 및 엡실론 값 설정

위험등급	동질그룹 크기 정규화	엡실론
매우 낮은 위험	0.8초과	1.0
낮은 위험	0.6초과 ~ 0.8이하	0.8
위험	0.4초과 ~ 0.6이하	0.6
높은 위험	0.2초과 ~ 0.4이하	0.4
매우 높은 위험	0.2 이하	0.2

2.3 차등 프라이버시를 적용한 데이터의 유용성 확인

데이터 프라이버시 역시 해결해야 할 중요한 문제이지만 본질적으로 데이터의 유용성을 과하게 훼손시키면 의미가 없다. 따라서 차등 프라이버시를 적용하여도 데이터의 유용성이 유지됨을 확인할 필요가 있다. 본 논문은 크게 기초 통계 비교분석 비교를 통해 데이터의 유용성을 증명하였다. 표 2는 원본 데이터의 기초 통계와 차등 프라이버시를 적용한 데이터의 기초 통계를 비교하였다.

(표 2) 원본 데이터 및 차등프라이버시 적용 데이터 통계 비교

통계	원본 데이터	dp 데이터
평균	54.539780	54.541620
표준편차	14.804764	15.285181
사분위수	55.0	54.57774

표 2를 통해 원본 데이터와 차등 프라이버시를 적용한 데이터의 기초 통계적 속성에서 큰 차이를 보이지 않아 데이터의 유용성을 유지할 수 있음을 알 수 있다. 이를 통해 프라이버시 보호와 데이터 분석의 효율성 간 균형을 달성하며, 실제 분석 결과가 데이터의 전반적인 패턴과 경향성을 유지하는 데 기여한다.

2.4 랜덤포레스트 모델의 적용

본 논문에서는 차등 프라이버시가 적용된 대출 가망 고객 데이터를 활용하여, 랜덤 포레스트 모델을 통해 대출 가망 고객을 예측하였다. 랜덤 포레스트 모델은 여러 개의 결정 트리(Decision Tree)를 결합한 앙상블 학습 알고리즘으로 다음과 같은 이유로 대출 고객 예측에 적합하다고 판단하였다.

첫째, 랜덤 포레스트 모델은 비선형 관계와 복잡한 상호작용을 잘 처리할 수 있다. 대출 승인 여부는 고객의 신용 점수, 소득, 채무 상태 등 여러 변수 간 복잡한 상호작용을 포함하며 변수의 중요도를 통해 어떤 변수가 중요한 역할을 하는지 쉽게 해석할 수 있다. 이를 통해 본 논문의 목적인 직업 특성 상 신용점수가 낮아 대출 접근이 어려우나 대출이 필요한 사람들을 정확히 식별할 수 있게 된다. 둘째, 기술적으로 랜덤 포레스트 모델은 과대적합을 방지하는 특성이 있다. 개별 결정 트리는 과대적합에 취약할 수 있으나 랜덤 포레스트는 다수의 트리를 훈련하며 그 결과를 평균화하기 때문에 모델이 안정적이며 성능이 뛰어나다. 즉, 훈련 데이터에 과도하게 맞추지 않아 새로운 데이터, 새로운 고객에 대한 예측 능력을 강화할 수 있다. 마지막으로 랜덤 포레스트는 노이즈에 강한 모델로, 불완전하거나 결측치가 일부 포함되어 있는 데이터셋에서도 안정적인 성능을 보여준다.

3. 결론

본 논문은 차등 프라이버시가 적용된 대출 가망 고객 데이터를 바탕으로, 랜덤 포레스트 모델을 활용하여 대출 가능성을 예측하는 모델을 제안한다. 해당 모델을 통해 대출 접근이 어려운 고객들을 신용 점수 등의 기존 평가 방식에서 벗어나 정밀하게 식별할 수 있으며 금융 포용을 강화하고자 하는 정책 목표를 실현시키고자 한다. 본 논문은 차등 프라이버시 기술을 적용함으로써 개인정보보호와 데이터의 유용성 사이의 균형을 유지할 수 있음을 확인하였다. 또한 랜덤 포레스트 모델의 비선형성과 복잡한 상호작용을 처리하는 능력을 통해 대출 예측 문제를 유용하게 해결할 수 있음을 입증하였다. 그러나 본 연구는 데이터 전처리 과정에서 비식별화 과정이 데이터를 일부 손실할 위험이 있다. 또한, 차등 프라이버시에 사용된 엡실론 값에 따라 개인정보 보호 수준이 달라지므로 다양한 엡실론 값에 따른 데이터 유용성 및 모델 성능의 차이를 더 정밀하게 분석할 필요가 있다. 향후 연구에서 차등 프라이버시의 엡실론 값을 조정하며 데이터 유용성과 데이터 보호의 최적 균형을 탐색하고 하이퍼 파라미터에 따른 모델의 성능을 분석하여 대출 예측의 정확도를 향상시키는 방향으로 연구를 진행할 수 있다. 이를 통해 금융기관이 정교한 대출 예측 모델을 통해 부실 대출을 방지하고 개인은 적절한 대출 기회를 제공받음으로써 금융 포용 정책의 효과를 극대화할 수 있을 것으로 기대된다.

감사의 글

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

참고 문헌

- [1] 고태수, 구본호. "차등 프라이버시(Differential Privacy)의 가능성과 한계." 서울대학교 인공지능정책 이니셔티브 이슈페이퍼, no. 03, 2020.
- [2] Apple. "Differential Privacy." Available: https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf. (Accessed: May 10, 2020).
- [3] 정수용, 홍도원, 서창호. "로컬 차분 프라이버시 실제 적용 사례연구: 프라이버시 보존형 설문조사." 정보보호학회논문지, 2020.
- [4] 김재환, 이민구. "의사 결정 트리와 랜덤 포레스트 모델 분석." 한국정보기술학회 학술대회자료, 2018. 6.

K-익명성 기반의 차분 프라이버시를 이용한 안전한 개인정보 활용

장승수
중앙대학교 산업보안학과
e-mail : dsml17@cau.ac.kr

Safe Utilization of Personal Information Using K-Anonymity Based Differential Privacy

Jang Seung Su
Dept. of Industrial Security, Chung-Ang University

1. 서론

빅데이터 시대가 도래하면서 방대한 양의 정보가 빠르게 수집되고 분석되는 것이 일상화되었다. 기업과 정부 기관은 이러한 데이터를 활용하여 새로운 가치를 창출하고, 이를 바탕으로 혁신적인 제품을 개발함과 동시에 맞춤형 서비스를 제공하고 있다. 그러나 데이터의 양과 중요성이 증가함에 따라 개인정보 보호의 필요성도 그 어느 때보다 중요해지고 있다. 특히, 개인의 민감한 정보가 무분별하게 수집되고 잘못 활용될 경우, 개인의 프라이버시가 심각하게 침해될 수 있으며, 이는 신뢰의 상실과 법적 문제를 초래할 수 있다. 빅데이터의 잠재력을 최대한 활용하면서도 개인정보 보호를 보장하는 것은 현대 사회가 직면한 중요한 과제 중 하나라고 할 수 있다. 그렇기에 데이터 분석과 처리 과정에서의 유용성 확보와 더불어 개인정보를 보호하기 위한 기술이 필요했고 그에 대한 요구로 PET(Privacy-Enhancing Technologies) 기술이 등장했다. PET 기술의 종류에는 대표적으로 ‘동형 암호’, ‘차분 프라이버시’, ‘합성데이터’ 등이 있다. 개인정보보호위원회에서는 개인정보보호 관련 각종 지침을 제공하고 있고 이 중에서는 PET 기술을 다루는 가이드라인도 존재한다. 2024년 5월에 공고된 ‘합성데이터 생성 참조모델’이 그 예시이다. 그러나 지금까지도 차분 프라이버시를 다루는 내용은 가명정보 처리 가이드라인의 부록에만 나타나 있으며 그마저도 실무적으로 적용해 볼 수 있는 내용이 아닌 이론적인 설명이 주를 이루고 있다. 차분 프라이버시는 데이터의 신뢰성을 유지하면서 개인의 프라이버시를 보호할 수 있도록 하는 중요한 기술 중 하나이다. 따라서 본 논문에서는 차분 프라이버시의 개념을 설명하고 실무 환경에서 차분 프라이버시를 어떻게 적용할 수 있는지에 대해 K-익명성 기반의 위험 평가에 따른 차등적 엡실론 적용 방식을 제시하고자 한다.

2. 원본데이터 기반으로 생성한 기존 합성데이터의 한계

개인화된 서비스를 제공하기 위해서는 방대한 양의 데이터를 수집하고 분석하는 과정이 필요하다. 이는 맞춤형 경험을

위해서는 사용자 개인의 데이터 공개가 필수적이라는 것을 의미한다. 그러나 여기에는 문제점이 분명 존재한다. 바로 개인의 데이터 중에서는 건강 정보, 주거지, 고유식별정보, 성향 등과 같은 민감 정보들도 포함되어 있다는 것이다. 따라서 민감 정보의 유출 문제를 방지하기 위해 원본데이터를 기반으로 생성한 합성데이터를 활용하여 데이터를 공개해 왔다. 다만, 이렇게 생성된 합성데이터도 유출 문제를 근본적으로 해결하기에는 한계가 있다. 그 이유는 다음과 같다.

공개된 합성데이터를 이용하여 특정 데이터로 특정 개인을 추론하는 추론(inference) 공격이 가능하다. 예를 들어 어떤 병원에서 의료 기록을 바탕으로 합성데이터를 공개했다고 가정해 보자. 의료 기록에는 나이, 질병, 성별, 치료 등의 민감 정보가 포함되어 있을 것이다. 물론 합성데이터이기엔 개별 환자 정보는 식별할 수 없도록 처리가 되어있지만, 원본데이터를 바탕으로 생성하였기에 통계적 특성은 여전히 유지된다. 만약 공격자가 특정 개인에 대한 사전 정보를 가지고 있고 합성데이터에서 이와 일치하는 정보가 매우 적다면 공격자는 해당 개인의 정보를 추론할 수 있게 된다. 혹은 데이터에 나타나는 특정한 패턴이나 통계적 특성을 분석하여 개인에 대한 추가 정보를 추론할 수도 있을 것이다. 마찬가지로 공개된 합성데이터 이외에도 여러 데이터와 연결하여 특정 개인을 추론하는 연계성(linkability) 공격도 가능하다[1]. 예시에서 알 수 있듯 원본데이터를 바탕으로 단순히 합성데이터를 생성하는 것만으로는 여전히 유출 문제가 존재한다. 그리고 이를 방지하기 위해 차분 프라이버시를 적용할 수 있다.

3. 차분 프라이버시와 엡실론

차분 프라이버시는 데이터 분석 중 개인의 프라이버시를 보호하기 위해 노이즈를 도입하는 기술이다. 특정 개인의 데이터가 포함되었는지 여부를 알아내기 어렵게 만드는 기술이다. 이 과정에서 중요한 역할을 하는 것이 바로 엡실론(ϵ)이라고 불리는 매개변수이다. 수학적으로 설명하자면 엡실론은 두 데이터베이스가 하나의 레코드만 다른 경우, 해당 데이터베이스에서 동일한 분석 쿼리 결과가 나올 확률이 얼마나 차이 나는지를 나타낸다[2].

앱실론은 차분 프라이버시가 어떤 식으로 적용될지를 결정하는 주요 요소이다. 앱실론 값이 작아질수록 개인정보 보호 정도가 더 강력해지고, 앱실론 값이 커질수록 보호 정도는 낮아지지만 데이터 유용성은 더 높아진다. 즉, 앱실론을 통해 프라이버시와 데이터 유용성 간의 균형을 맞출 수 있다. 앱실론에 대한 설명에서 알 수 있듯이 차분 프라이버시는 개인정보 보호와 더불어 데이터 유용성 유지, 수학적 신뢰성 보장, 대규모 데이터에의 광범위한 적용 가능성이라는 장점이 존재한다. 반면, 소규모 데이터에서의 노이즈 추가로 인한 데이터 정확도 저하, 노이즈를 추가하고 처리하는 과정에서 발생하는 비용 이슈, 차분 프라이버시 알고리즘을 구현하기 위한 높은 수준의 기술적 지식과 경험 요구, 정책적 논의 필요 등의 제한 사항도 존재하기 때문에 이 역시도 차분 프라이버시를 적용하기 위해 충분히 고려되어야 한다.

4. 처리목적에 따른 앱실론 범위 결정

그렇다면 실무 환경에서는 차분 프라이버시를 어떻게 적용할 수 있을까? 계속해서 강조하지만 차분 프라이버시를 적용하는 목적은 개인정보 보호와 더불어 데이터 유용성도 확보하는 것이기에 결국 앱실론 값을 상황에 맞게 적절히 선택하는 것이 차분 프라이버시 적용의 핵심이 되겠다. 앱실론 값은 데이터를 처리하는 목적에 따라 수치가 달라져야 한다. 만약 처리 목적이 공개용이라면 개인정보 보호가 엄격히 적용되어야 한다. 그런 경우에는 앱실론 값을 0.1~0.5 사이로 정함으로써 데이터 변조를 다소 크게 적용한다. 반면 처리 목적이 데이터 분석용이라면 내부 분석용으로 데이터 유용성이 높아야 하므로 앱실론 값을 5.0~10.0 사이로 정함으로써 데이터 변조를 작게 한다. 일차적으로 처리 목적에 따라 앱실론 값의 범위를 정했다면 이제는 정한 범위 내에서 앱실론을 데이터에 적용해야 한다. 동일한 앱실론 값을 일괄적으로 적용하는 방법도 생각해 볼 수는 있겠지만 그렇게 된다면 전체 데이터에 동일한 수준의 개인정보 보호가 적용된다. 이는 데이터의 특성을 전혀 고려하지 않은 방법으로 일부 데이터는 필요 이상의 개인정보 보호를 받을 것이며, 또 다른 데이터는 충분한 수준의 개인정보 보호를 받지 못할 가능성이 존재한다. 따라서 앱실론 값을 설정한 범위 내에서 차등적으로 적용하는 방법을 추천하며 본 논문은 K-익명성 기반의 위험 평가에 따른 차등적 앱실론 적용 방식을 제시한다.

(표 1) 처리목적에 따른 차등적 앱실론 범위 예시

처리목적	개인정보 보호와 데이터 유용성 관계	Differential Privacy 앱실론 범위
공개용	개인정보 보호 수준이 엄격히 적용되어야 함	0.1~0.5
교육용	내부 교육용으로 처리환경에 따라 개인정보 보호와 데이터 유용성 적용 가능	0.5~1.0
테스트용	내부 기술검증용으로 처리환경에 따라 개인정보 보호와 데이터 유용성 적용 가능	1.0~5.0
데이터 분석용	내부 분석용으로 데이터 유용성이 높아야 함	5.0~10.0

5. K-익명성 기반 위험 평가에 따른 차등적 앱실론 적용 방식의 이점과 한계

K-익명성은 특정 데이터가 다른 개인과 구별되지 않게 만드는 것을 목표로 하며, 동일한 속성을 가진 최소 K명의 각기 다른 사람이 있는 경우 해당 데이터는 재식별이 어렵다고 간주한다. 그렇기에 K-익명성을 사용하여 각각의 데이터의 재식별 위험을 평가한 후, 높은 위험도를 보이는 데이터에는 더 작은 앱실론 값을, 위험도가 낮은 데이터에는 더 큰 앱실론 값을 적용한다. 이렇게 한다면 재식별 위험이 낮은 데이터의 유용성을 높이면서도 재식별 위험이 높은 데이터는 민감도를 낮춤으로써 보다 엄격하게 개인정보를 보호할 수 있다는 장점이 있다. 또한 재식별 공격을 수행하는 공격자가 데이터별 개인정보 보호 수준을 예측하기가 까다롭기에 공격 저항성이 강화된다는 이점도 존재한다. 다만, K-익명성 기반의 위험 평가에 따른 차등적 앱실론 적용 방식도 역시 일부 제한 사항을 가지고 있다. 첫째, K-익명성의 내재적인 한계로 인해 외부 데이터와 결합할 경우에는 프라이버시가 취약해질 수 있다. 둘째, 값을 동적으로 할당하는 과정에서 추가적인 계산 리소스가 필요할 수 있으며, 이에 따라 시스템의 복잡성이 증가할 가능성이 있다. 따라서 실무에서 해당 방식을 적용할 때는 이점뿐만 아니라 제한 사항까지 모두 고려하여 적절한 개인정보 보호 조치가 이루어져야 한다.

6. 결론

본 논문에서는 차분 프라이버시의 개념과 K-익명성을 기반으로 실무에서는 어떻게 적용할 수 있을지 다루었다. 차분 프라이버시의 핵심은 적절한 앱실론 값을 통해 개인정보 보호와 데이터 유용성 간의 균형을 맞추는 것이다. 필자가 제시한 K-익명성 기반의 차등적 앱실론 적용 방식을 통해 데이터 항목별 맞춤형 개인정보 보호를 적용함으로써 재식별 위험을 줄이면서도 데이터 유용성을 극대화할 수 있다. 다만, 앞서 언급했듯이 K-익명성의 한계와 시스템 복잡성 등 실무에서 고려해야 할 제한 사항도 존재하기 때문에 충분히 고려하여 차분 프라이버시를 적용해야 할 것이다. 차분 프라이버시를 더 효과적으로 적용하기 위해서는 기술적 부분과 더불어 정책적 논의 역시 필요하다. 이외에도 앞으로 차분 프라이버시와 연계한 다양한 기술들이 산업 분야에서 어떻게 활용될 수 있을지에 대한 구체적인 연구도 필요할 것이다. 이러한 노력을 통해 빅데이터 시대에 데이터 유용성과 개인정보 보호라는 두 가지 목표를 모두 달성할 수 있는 계기를 마련할 수 있을 것이다.

감사의 글

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

참고 문헌

- [1] 강준영. "차분 프라이버시를 만족하는 제한 데이터 생성을 위한 딥러닝 모델 활용 기술 연구." 공주대학교 석사학위논문, 2021.
- [2] 정강수, 박석. "차분 프라이버시 기반 비식별화 기술에 대한 연구." 정보보호학회지, 제28권 제2호, pp. 61-77, 2018.

GAN 기반 데이터 증강을 활용한 개인정보 보호 강화 및 멤버십 추론 공격 방어 기술 연구

전하영
 중앙대학교 산업보안학과
 e-mail : lipun0258@cau.ac.kr

Research on privacy enhancement and defense against membership inference attacks using GAN-based data augmentation

Jeon Ha Young
 Dept. of Housing & Interior Design, Chung-Ang University

1. 서론

인공지능(AI)과 머신러닝 기술의 발전은 다양한 산업에서 혁신을 가져왔지만, 민감한 데이터를 포함한 모델 학습이 증가함에 따라 개인정보 보호 문제가 대두되고 있다. 특히, 멤버십 추론 공격(Membership Inference Attack)은 공격자가 모델이 학습한 데이터 중 특정 데이터의 포함 여부를 추론할 수 있어, 개인 정보 유출의 위험이 크다. 멤버십 추론 공격은 의료 및 금융 데이터와 같은 민감한 정보를 다룰 때 심각한 위협이 된다. 이에 따라, 본 연구는 GAN(Generative Adversarial Networks)을 활용하여 데이터를 증강함으로써, 멤버십 추론 공격에 대한 방어 성능을 높이는 방법을 제시한다. 데이터 증강을 통해 모델의 학습 성능을 유지하면서도 개인정보를 보호할 수 있는 가능성을 탐구한다.

2. 연구 배경 및 기법

2.1 멤버십 추론 공격

멤버십 추론 공격(Membership Inference Attack)은 머신러닝 모델이 학습한 데이터에 특정 데이터가 포함되었는지를 공격자가 추론할 수 있는 공격 방식이다. 이러한 공격은 특히 민감한 정보가 포함된 데이터셋에서 심각한 문제를 일으킬 수 있으며, 개인 정보 유출의 위험성을 높인다. 예를 들어, 병원에서 사용되는 의료 데이터나 금융 기관의 신용평가 모델에서 멤버십 추론 공격이 성공할 경우, 특정 개인의 민감한 정보가 유출될 수 있다.

2.2 프라이버시 공격

프라이버시 공격(Privacy Attack)은 멤버십 추론 공격을 포함한 다양한 공격 방식으로, 모델이 학습한 데이터에서 개별 사용자의 정보를 추출하려는 시도를 의미한다. 프라이버시 공격은 주로 데이터가 풍부한 모델에서 발생하며, 공격자가 모델의 결과를 분석하여 민감한 데이터를 유출하려고 할 때 발생한다. 기존의 차등 프라이버시(Differential Privacy)는 이러한 공격을 방어하기 위한 방법으로, 데이터

에 노이즈를 추가하여 개별 데이터를 은폐하는 역할을 하지만, 성능 저하 문제를 야기할 수 있으며, 이를 보완하기 위해 GAN 기반 데이터 증강 기법이 제안되었으며, 이를 통해 성능 저하 없이 공격을 방어할 수 있다.

2.3 GAN을 활용한 데이터 증강

본 연구는 차등 프라이버시의 한계를 보완하고, 멤버십 추론 공격을 방어하기 위해 GAN 기반 데이터 증강 기법을 제안한다. WGAN-GP(Wasserstein GAN with Gradient Penalty)와 SNGAN(Spectral Normalization GAN)을 사용하여 데이터를 증강함으로써, 원본 데이터에 노이즈를 추가하지 않고도 민감한 정보를 보호할 수 있다. 이러한 데이터 증강을 통해 모델의 학습 성능을 유지하면서도 개인정보 보호를 강화할 수 있다.

3. 실험 설계 및 결과

본 연구에서는 GAN을 활용한 데이터 증강이 멤버십 추론 공격에 대한 방어 성능을 어떻게 향상시키는지 평가하기 위해 실험을 설계하였다. 실험은 의료 데이터를 기반으로 진행되었으며, WGAN-GP와 SNGAN을 활용하여 증강된 데이터와 원본 데이터를 비교하였다. 각 모델에 대해 멤버십 추론 공격을 수행한 후, 성능 및 공격 방어 효과를 측정하였다.

3.1 실험 설정

- **데이터 셋:** 의료 데이터를 사용하였으며, 민감한 개인 정보가 포함된 항목을 대상으로 증강을 진행하였다.
- **모델:** Random Forest 모델을 학습에 사용하였으며, 원본 데이터와 GAN으로 증강된 데이터를 각각 학습시켰다.
- **학습 시간:** 각 GAN 모델의 학습 시간은 평균 3시간 소요되었으며, 증강된 데이터를 포함한 전체 모델 학습 시간은 6시간이었다.
- **평가 지표:** 정확도(Accuracy), 정밀도(Precision), 재현율

(Recall), F1 Score를 사용하여 모델 성능을 평가하였고, 멤버십 추론 공격 성공률을 측정하였다.

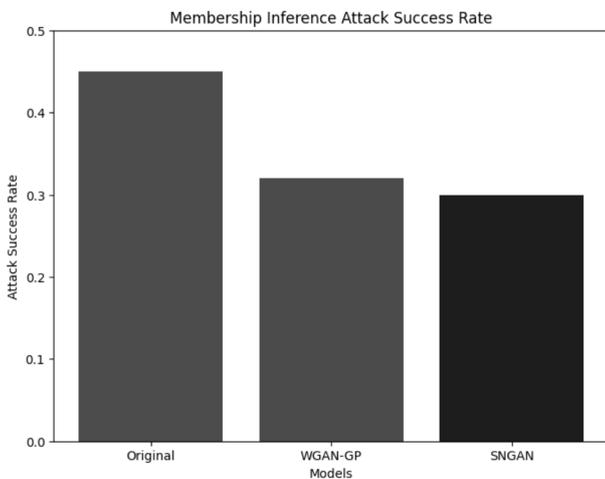
3.2 실험 결과

WGAN-GP로 증강된 데이터 셋을 학습한 모델은 멤버십 추론 공격 성공률이 0.32로, 원본 데이터만 학습한 모델(성공률 0.45)보다 낮은 결과를 보였다. SNGAN으로 증강된 데이터 셋을 학습한 모델은 성공률이 0.30으로 가장 낮았으며, 이는 데이터 증강이 공격 방어에 효과적임을 입증한다. SNGAN 모델이 더 나은 성능을 보인 이유는 데이터 증강 과정에서 안정성을 제공하기 때문이다.

(표 1) 실험 결과 비교 표

모델	공격 성공률	정밀도	재현율	F1 Score
원본 데이터	0.45	0.72	0.68	0.70
WGAN-GP	0.32	0.80	0.70	0.74
SNGAN	0.30	0.82	0.72	0.76

표1은 원본 데이터와 증강된 데이터에 대해 멤버십 추론 공격 성공률과 모델 성능을 비교한 결과를 보여준다. SNGAN 모델이 가장 낮은 공격 성공률을 기록하였으며, 이는 GAN을 통한 데이터 증강이 개인정보 보호에 효과적임을 나타낸다.



(그림1) 실험 결과 그래프

이 그래프는 각 모델에 대해 멤버십 추론 공격 성공률을 시각적으로 보여준다. SNGAN 모델에서 공격 성공률이 가장 낮게 나타나며, 데이터 증강 기법이 방어에 유리함을 보여준다.

4. 결론

본 연구에서는 GAN(Generative Adversarial Networks)을 활용한 데이터 증강 기법이 멤버십 추론 공격에 대한 방어에 효과적임을 확인하였다. WGAN-GP와 SNGAN을 통해 생성된 데이터 셋을 학습한 모델은, 원본 데이터만을 학

습한 모델에 비해 공격 성공률이 현저히 낮았다. 이는 데이터 증강이 민감한 정보의 노출을 줄이는 동시에 모델의 성능을 유지하는 데 기여할 수 있음을 시사한다.

특히, SNGAN을 활용한 증강 데이터가 가장 낮은 멤버십 추론 공격 성공률을 기록했으며, 이는 데이터의 다양성을 증대시키면서도 원본 데이터에 대한 의존성을 줄일 수 있음을 보여준다. 본 연구는 GAN을 활용하여 멤버십 추론 공격에 대한 방어 기법을 실험적으로 입증하였으며, 기존의 차등 프라이버시 기법과 비교했을 때 성능 저하 없이 개인정보 보호를 강화할 수 있는 가능성을 제시하였다.

향후 연구에서는 실시간 AI 기반 프라이버시 보호 기법과의 결합을 통해 멤버십 추론 공격뿐만 아니라 다른 유형의 개인정보 침해 공격에 대한 방어력을 강화하는 방안도 탐구할 필요가 있다.

감사의 글

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

참고문헌

- [1] Cynthia Dwork. "Differential Privacy." Lecture Notes in Computer Science, vol. 4052, pp. 1-12, 2006.
- [2] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Networks." Advances in Neural Information Processing Systems, vol. 27, 2014.
- [3] Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. "Deep Learning with Differential Privacy." Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 308-318, 2016.
- [4] Ilya Mironov. "Rényi Differential Privacy." 2017 IEEE 30th Computer Security Foundations Symposium (CSF), pp. 263-275, 2017.
- [5] Arjun Nitin Bhagoji, Supriyo Chakraborty, Prateek Mittal, and Seraphin Calo. "Analyzing Federated Learning through an Adversarial Lens." Proceedings of the 36th International Conference on Machine Learning, pp. 634-643, 2019.

차등 프라이버시와 데이터 합성 기술을 적용한 신경망 모델의 성능 및 프라이버시 보호 분석

조성원
중앙대학교 산업보안학과
e-mail : won66399@cau.ac.kr

Analyzing the Performance and Privacy Protection of Neural Networks with Differential Privacy and Data Synthesis Techniques

Seong Won Cho
Dept. of Industrial Security, Chung-Ang University

요 약

본 연구는 데이터 거래가 유형 자산에서 무형 자산으로 전환되면서 데이터 프라이버시 보호의 중요성이 증대되는 상황에서, 차등 프라이버시(DP)와 데이터 합성 기술이 신경망 모델의 성능 및 프라이버시 보호에 미치는 영향을 분석한다. 2020년도 데이터 3법 개정으로 가명처리 데이터 거래가 중요해짐에 따라, 본 논문은 두 가지 실험을 수행한다.

첫째, 로지스틱 회귀와 의사결정 나무 모델에 차등 프라이버시를 적용해 epsilon 값에 따른 모델 성능과 프라이버시 보호 간의 관계를 분석한다. 둘째, CNN 모델과 MNIST 데이터셋을 사용해 데이터 합성 기술을 적용하고, 합성 데이터가 모델의 정확도 및 프라이버시 보호 수준에 미치는 영향을 평가한다.

이 연구는 프라이버시 보호 기술이 AI 모델 성능에 미치는 영향을 종합적으로 분석하며, 최종적으로 LLM과 차등 프라이버시의 관계를 파악하는 것을 목표로 한다. 데이터 거래의 새로운 패러다임에서 실용적인 프라이버시 보호 기법을 제시하는 데 기여한다.

Acknowledgement

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

딥페이크 범죄 경로 분석을 통한 다차원적 대응 방안 제언

최강규
중앙대학교 산업보안학과
e-mail : racoon3719@cau.ac.kr

Deep Fake Crime Path Analysis multi-dimensional response

Kang Kyu Choi
Chung-Ang University Industrial Security

요 약

인공지능 기술의 발달로 딥페이크 범죄는 점점 정교해지고 있으며, 특히 프라이버시 침해와 관련된 문제가 심각해지고 있다. 이를 막기 위해 딥페이크 콘텐츠에 워터마크를 삽입하는 기술적 방어책이 도입되고 있지만, 이 방법은 제거되거나 우회될 가능성이 있어 프라이버시 침해를 완전히 차단하기에는 한계가 있다. 특히, 소셜 미디어, 암호화된 플랫폼, 다크웹 등에서 딥페이크의 유통이 활발히 이루어지면서 기존의 기술적 대응만으로는 범죄 경로를 차단하기 어려운 상황이다.

본 연구는 딥페이크 범죄가 어떻게 발생하고 유통되는지 분석함으로써, 프라이버시 침해가 일어나는 구체적인 경로를 파악하고자 한다. 이를 바탕으로 기술적 방어의 한계를 지적하고, 법적 규제 강화, 국제 협력, 사회적 인식 제고 등 다방면에서 프라이버시 보호를 실현할 수 있는 종합적인 정책 전략을 모색한다. 궁극적으로는 딥페이크 범죄로 인한 프라이버시 침해를 효과적으로 방지할 수 있는 정책적 대응 방안을 제시하는 것을 목표로 한다.

Acknowledgement

이 성과(논문)는 정부(교육부)의 지원을 받아 수행된 연구임 (2024년 부처 협업형 인재양성사업[정보보안 분야], No. 2024 개인정보보호-006)

여행조각가: LLM 기반 여행 영상 스토리보드 제작 서비스

김도균, 조민규, 박상오
중앙대학교 소프트웨어학부

e-mail : ehrbs7181@gmail.com, mgjo@cslab.cau.ac.kr, sopark@cau.ac.kr

Travel Sculptor: LLM Based Travel Video Storyboard Generation Service

Dogyun Kim, Mingyu Jo, Sangoh Park
School of Computer Science and Engineering, Chung-Ang University

요 약

코로나19 팬데믹 이후 여행 관련 콘텐츠에 대한 수요가 급증했으며 개인의 여행 빈도 또한 크게 증가했다. 중앙대학교 학생들을 대상으로 실시한 설문조사 결과, 많은 학생들이 자신의 여행 경험을 영상으로 기록하여 추억하고자 하는 수요가 있음이 확인되었다. 그러나 전문적인 지식이나 경험이 없는 일반인들이 효과적인 여행 영상 스토리보드를 구성하는 것은 여전히 어려운 과제로 남아있다. 본 논문에서는 Large Language Model(LLM)을 활용하여 사용자가 제공한 여행지와 동행자 정보를 기반으로 자동으로 스토리보드를 생성해주는 서비스인 여행조각가를 제안한다. 여행조각가는 사용자 입력 데이터를 심층적으로 분석하여 개인화된 콘텐츠를 제작하도록 도와준다. 사용자 입력 데이터를 분석하여 다양한 여행 콘셉트에 맞는 영상 타이틀, 인트로 및 아웃트로 옵션을 제시하고, 이를 종합하여 최종 스토리보드를 생성함으로써, 전체적인 영상 제작 과정을 효율적으로 지원한다. 또한 Youtube Data API와 한국관광공사에서 제공하는 데이터를 활용하여 최신 트렌드에 적합한 스토리보드를 제공할 수 있도록 한다. 제안하는 여행조각가 서비스의 의의는 여행 영상 제작의 진입 장벽을 낮추어 초보자들도 고품질의 여행 영상을 제작할 수 있도록 지원하는 데 있다. 또한, 영상 편집 초보자들에게 효과적인 기획 도구를 제공함으로써 제작 시간을 단축하고 더 효율적인 여행 이야기 전달이 가능하게 한다. 이를 통해 여행 콘텐츠 제작의 대중화와 질적 향상에 기여할 것으로 기대된다. 본 시스템의 효과성은 실제 사용자 테스트를 통해 정성적으로 검증될 것이다. 제공된 입력 데이터를 바탕으로 생성된 스토리보드에 대한 80% 이상의 높은 만족도와 영상 제작 시간이 30% 이상 단축될 것을 기대한다.

Char각: 머신러닝 기반 촬영 가이드 플랫폼

정현서, 김경연, 김동현, 박상오
중앙대학교 소프트웨어학부

e-mail : hyunseo0412@naver.com, ky0207@cau.ac.kr, dhkim@cslab.cau.ac.kr,
sopark@cau.ac.kr

CharKak: Machine Learning-Based Photography Guide Platform

Hyunseo Jung, Kyeongyeon Kim, Donghyeon Kim, Sangoh Park
School of Computer Science and Engineering, Chung-Ang University

요 약

스마트폰 카메라의 발달과 SNS의 성장으로 인해 사진 촬영 기술의 중요성이 크게 증가하고 있다. 많은 사람들이 사진을 더 잘 찍고 싶어 하지만, 촬영 기술을 체계적으로 향상시키는 방법이 부족하여 어려움을 겪고 있다. 기존의 사진 관련 앱은 주로 이미지 수정 및 편집에만 초점을 맞추고 있어, 사진을 잘 찍는 기술 학습에 도움을 주는 서비스가 부족하다. 이에 따라, 사진 촬영 경험이 적거나 촬영 기술을 향상시키고자 하는 사용자들을 위한 교육 플랫폼의 필요성이 대두되고 있다. 본 논문에서는 사용자가 찍은 사진을 AI로 분석하여 피드백을 제공하는 웹 기반 플랫폼인 Char각을 제안한다. Char각은 사용자가 찍은 사진을 업로드하면 사진을 분석하여 사용자에게 필요한 사진 촬영 기술 학습 자료와 점수를 제공하며, 사용자 사진 기술에 대한 성장 그래프 제공을 통해 사용자의 촬영 기술 향상에 대한 동기를 유발한다. 또한 사진을 찍을 수 있는 가상환경을 제공함으로써 사용자가 촬영에 익숙해질 수 있도록 유도한다. Char각은 React, SpringBoot, MySQL, PyTorch 및 OpenCV를 사용하여 구현되었다. 설문조사를 통한 Char각의 평가 결과, AI 모델의 사진 분석은 전문가 평가와 92% 일치하는 결과를 보여주었다. 평균 처리 시간은 3초로, 실시간 피드백 제공에 충분한 성능을 나타냈다. 85%의 사용자가 피드백이 촬영 기술 향상에 도움이 된다고 응답하였으며, 가상 카메라와 학습 자료 제공 기능이 특히 높은 평가를 받았다. Char각은 사용자가 사진 촬영 기술을 체계적으로 향상시킬 수 있는 맞춤형 학습 도구를 제공하며, AI 기반 피드백을 통해 사진 품질을 개선하는 데 도움을 줄 것이다. 또한, 가상 촬영 기능을 통해 다양한 촬영 경험을 제공하여 실질적인 기술 향상을 기대할 수 있다.

게이트웨이: 웹2와 웹3의 연결

추교현¹⁾, 조민규²⁾, 박상오³⁾

¹⁾ 중앙대학교 기계공학부, 소프트웨어벤처융합전공

^{2,3)} 중앙대학교 소프트웨어학부

e-mail : ckh0601@cau.ac.kr, mgjo@cslab.cau.ac.kr, sopark@cau.ac.kr

Gateway: Connection between Web2 and Web3

Kyohyun Choo¹⁾, Mingyu Jo²⁾, Sangoh Park³⁾

¹⁾ Mechanical Engineering and Software Venture, Chung-Ang University

^{2,3)} School of Computer Science and Engineering, Chung-Ang University

요 약

Web2 개발자들에 대한 공정한 보상 부재는 디지털 생태계의 발전을 저해하는 문제 중 하나다. Web2는 중앙화된 서버 기반의 애플리케이션 개발에 중점을 두고 있으며, Web3는 탈중앙화와 투명성을 기반으로 하는 보상 시스템을 제공한다. 하지만, Web2 개발자들은 Web3의 인프라와 보상 시스템을 쉽게 활용할 수 없으며, 중앙화된 플랫폼에서 보상 구조는 불투명하거나 제한적이다. 이러한 문제를 해결하기 위해 Web2 애플리케이션이 Web3 보상 시스템과 연결되어 개발자들이 공정한 보상을 받을 수 있는 Gateway 플랫폼을 제안한다. Gateway 플랫폼은 세 가지 핵심 컴포넌트로 구성된다: 1) Web2 인터페이스, 2) CosmWasm 기반 스마트 컨트랙트, 3) 오라클 시스템. Web2 인터페이스는 기존 Web2 개발자들이 별도의 Web3 지식 없이도 서비스를 등록하고 트랜잭션을 발생시킬 수 있게 한다. CosmWasm 기반 스마트 컨트랙트는 Web2에서 발생한 트랜잭션을 Archway 블록체인상에서 처리하며, Gateway Bot이라는 오라클 시스템은 Web2와 Web3 간의 데이터 및 이벤트를 실시간으로 동기화한다. Web3 보상 시스템으로는 Archway 블록체인을 활용한다. Archway는 탈중앙화 애플리케이션(dApp)에서 발생하는 트랜잭션의 처리량에 비례한 ARCH 토큰으로 해당 dApp 개발자들에게 인센티브를 제공한다. Gateway를 통해 등록된 Web2 애플리케이션의 활동은 Archway의 탈중앙화된 블록체인 노드들에 의해 투명하게 검증되고 기록된다. 이를 통해 Web2 개발자들은 애플리케이션에서 발생하는 트랜잭션 수에 비례한 ARCH 토큰을 보상으로 받을 수 있다. 또한, Gateway는 EVM(Ethereum Virtual Machine) 기반 블록체인과의 호환성도 제공하도록 설계되어 다양한 블록체인 네트워크와 상호작용할 수 있는 확장성을 갖추고 있다. Web2 활동을 Web3 인센티브 시스템에 통합하여 투명하고 공정한 보상 구조를 제공함으로써 Gateway는 Web2와 Web3 생태계 간의 기술적, 경제적 격차를 줄일 수 있게 된다. 이러한 접근은 Web2 개발자들에게 Web3의 인프라를 활용하여 공정한 보상을 받을 수 있는 기회를 창출한다. 이는 개발자들이 더 나은 애플리케이션을 개발하도록 유도하고, 생태계의 지속 가능한 성장을 도모하는 중요한 역할을 하며, 탈중앙화된 인터넷 경제의 발전을 가속하는 의의를 갖는다.

Art Bidding Adventure

아트비딩어드벤처

학술대회 참가자에게 드리는 특별한 행운!

행운을 잡아라

최승희 작가



A. RE-ACT No.11



류시천 작가



B. Dreaming of Laputa



최창훈(단월) 작가



C. 내 마음을 받아줘



하나! 오전10:00(10월12일)에 개최식장(SKT홀)으로 온다.

둘! QR코드에 접속한다.

셋! 접속사이트에서 원하는 숫자 세자리(예: 025)를 입력한다.

넷! 당첨되면 작가에게 직접 작품을 선물로 받는다.

다섯! 사무실 벽면에 걸어 놓고 친구들에게 자랑한다.

여섯! 다음 학술대회에도 참여해서 더 큰 행운을 잡는다.

후원

