



한국정보과학회
Korean Institute of Information Scientists and Engineers

제 27 권 제 1 호
Vol. 27 No. 1



2025

제 27 회

한국 소프트웨어공학 학술대회 논문집

Proceedings of the 27th Korea Conference on
Software Engineering (KCSE 2025)

- 일시: 2025년 1월 20일(월) ~ 1월 22일(수)
- 장소: 강원도 평창 한화리조트(휘닉스파크점)

주최: 한국정보과학회, 한국정보처리학회

주관: 한국정보과학회 소프트웨어공학 소사이어티
한국정보처리학회 소프트웨어공학연구회

후원:  **SOLUTIONLINK**

(주)비트컴퓨터

(주)포멀웍스, (주)다한테크, (주)모아소프트

브이플러스랩(주), 슈어소프트테크(주), (주)이에스지(ESG)

한국정보통신기술협회(TTA)

세션 E1. 요구사항 분석

- 2025년 1월 22일 (수) 오전 10:40-11:50 / 그랜드홀 2
- 좌장: 이선아 교수(경상대)

드론 충돌 회피에 대한 초기 안전성 분석 기법에서의 GPT-4o 활용 연구

김태환, 이정원, 이선아 (경상국립대)

건전성 테스트를 위한 공간 및 운용 요구사항 분석 및 테스트 프로그램 설계 방법

양희찬, 최민서, 김진세, 이정원 (아주대)

Software Process 내에서 AI 지향 요구 공학 적용한 사례 (단편논문)

진예진, 서채연 (홍익대), 공지훈 (투스퀘어), 김기두 (한국정보통신기술협회), 김영철 (홍익대)

복잡한 문맥 개선을 위한 C3Tree 모델 기반 F1-Score 개선 사례 (단편논문)

김장환, 양진모 (홍익대), 장우성 (넥스트솔), 김영철 (홍익대)

[초청발표] PBE-based Selective Abstraction and Refinement for Efficient Property Falsification of Embedded Software (FSE 2024), 김요엘, 최윤자(경북대), The ACM International Conference on the Foundations of Software Engineering (FSE 2024) 국제학술대회 발표 논문

발표자: 김요엘 (경북대)

세션 E2. 지능형 SW 품질

- 2025년 1월 22일 (수) 오전 10:40-11:50 / 세미나실1
- 좌장: 백종문 교수(KAIST)

데이터 포맷에 따른 RAG 성능 비교와 LLM 기반 평가 체계의 실증적 분석 (단편논문)

김민지 (편진), 강승식 (국민대)

문맥 상호 연관 가중치의 중요성 : 구조화된 데이터에서 RAG 시스템의 환각 현상 완화를 위한 최적화 연구 (산업체논문)

이재문, 김예은, 김현지((주) 시즌 연구팀)

ISO/IEC TS 4213 AI 성능 측정 표준을 이용한 비교 숙련도 프로그램 개발에 관한 연구 (단편논문)

황현후, 김광훈, 김한결 (한국정보통신기술협회)

도메인 특화 LLM 모델의 프롬프트 최적화를 위한 에이전트형 프롬프트 (단편논문)

이지웅, 이철웅 (고려대)

딥페이크 탐지를 위한 지식 증류 기반 경량화 모델 (학부생논문)

이지수, 김지우, 김준화 (건양대)

복잡한 문맥 개선을 위한 C3Tree 모델 기반

F1-Score 개선 사례

김장환¹, 양진모², 장우성³, 김영철^{4*}
홍익대학교 소프트웨어공학연구실^{1,2,4*}
넥스트솔³

lentoconstante@hongik.ac.kr¹, yjmd2222@g.hongik.ac.kr²,
jwshurray@gmail.com³, bob@hongik.ac.kr^{4*}

Best Practices on Enhancing F1-Score

Based on C3Tree Model for complex Sentence Refinement

Janghwan Kim¹, Jinmo Yang², Woo Sung Jang³, R. Young Chul Kim^{4*}
Software Engineering Laboratory, Hongik University^{1,2,4*}
NextSol³

요약

현재, 다양한 인공지능 기반 도구들이 증가한다. 특히, 음성인식, 텍스트 분석, 감성분석 등 자연어 처리 분야의 관심이 증가하고 있다. 하지만, 한국어 자연어 처리 학습모델들은 대부분 정제되지 않은 초거대 데이터로 학습이 되고 있다. 한국어의 특성상 학습모델의 정확도나 성능 등 제한점이 존재한다. 이를 개선하기 위해, 한국어 자연어 처리 모델의 F1-Score 향상을 위해 C3Tree 모델 기반 텍스트 분류 방법을 활용하여 데이터를 정제한다. 이 정제된 텍스트 데이터는 소규모 데이터 환경에서도 안정적 학습을 위해 랜덤 포레스트(Random Forest) 기법으로 학습한다. 이를 통해, 한국어의 복잡성을 효과적으로 처리한다, 향후, 자연어 기반 3D 저작 환경과 같은 멀티모달 인터페이스 제어 기술에 활용될 수 있기를 기대한다. 또한, 한국어로 다양한 도메인에 적용할 수 있기를 기대한다.

Abstract

Currently, many AI-based tools are being developed. Natural Language Processing (NLP) fields like speech recognition, text analysis, and sentiment analysis are gaining more attention. However, most Korean NLP models are trained with large-scale data. Due to the characteristics of the Korean language, these models have limitations in accuracy and performance. To improve this, we use a C3Tree-based text classification method to refine data and enhance the F1-score of Korean NLP models. The refined data is trained with a Random Forest method, which ensures stable learning even in small data environments. This approach effectively handles the complexity of the Korean language. In the future, this method could be applied to multimodal interface technologies, such as natural language-based 3D authoring environments. It is also expected to support various domains in Korean.

1. 서론

최근 인공지능 기술의 급속한 발전과 함께, 3D 콘텐츠 제작 및 활용이 다양한 산업 분야에서 주목받고 있다. 애니메이션, 게임, 웹툰과 같은 콘텐츠 산업은 물론, 메타버스와 같은 차세대 디지털 환경에서도 3D 객체와 장면 저작 기술은 핵심적인 역할을 하고 있다[1]. 문제점은 현존하는 3D 제작 기술은 여전히 전문가의 수작업에 크게 의존하고 있으며, 이는 제작 공정의 효율성을 저해하는 주요 요인으로 작용한다. 특히, 스토리텔링과 장면 저작을 위한 자연어 학습 데이터의 부족은 비전문가가 쉽게 접근할 수 있는 3D 모델링 기술 개발의 큰 장애물이다[1].

현재 메타(Meta)의 ‘빌더봇’과 엔비디아(NVIDIA)의 ‘갠버스 3D’와 같은 기술은 음성이나 단일 이미지를 기반으로 3D 환경을 생성할 가능성을 제시하고 있다.

그러나, 이러한 기술은 특정한 상황에서만 효과적으로 작동하며, 다양한 스토리와 장면 맥락을 반영한 3D 제작을 위한 저작 도구 개발은 여전히 도전적인 과제로 남아 있다. 이를 해결하기 위해 멀티모달 인터페이스를 활용한 기술 제어 및 개선 방안에 대한 연구가 활발히 진행되고 있다. 특히, 멀티모달 제어에 필요한 자연어 처리 기술은 빠르게 발전하고 있지만, 한국어의 특성을 고려한 문어체와 구어체의 차이, 문장의 모호성과 유사성을 효과적으로 처리하는 데는 여전히 한계가 존재한다. 이러한 문제는 한국어 기반 자연어 처리 모델이 3D 저작도구 제작을 지원하기 위한 핵심 데이터를 충분히 갖추지 못했기 때문이기도 하다. 따라서, 한국어 기반 대화형 멀티모달 인터페이스를 통해 사용자 의도를 정확히 반영하고, 지능형 검색 및 에셋 추천 기술을 활용하여 3D 장면을 신속하게 구성할 수 있는 통합적 기술 개발이 필요하다[2].