

생성형 AI 저작권 침해 대응을 위한 LangGraph 기반 LLM 에이전트 설계

정현우, 최윤서, 최동훈, 손경아*
아주대학교

jhw0714@ajou.ac.kr, chldbstj0512@ajou.ac.kr, dhc4003@ajou.ac.kr, *kasohn@ajou.ac.kr

LangGraph-based LLM Agent Design for Addressing Copyright Infringement by Generative AI

Hyun Woo Jeong, Yun Seo Choi, Dong Hun Choi, Kyung Ah Sohn*
Ajou Univ.

요 약

본 연구는 생성형 AI 기반 이미지의 저작권 침해 여부를 기술적으로 판별하고, 법률적 자문을 제공하는 LangGraph 기반 LLM 에이전트를 제안한다. 생성형 인공지능의 발전은 다양한 분야의 콘텐츠 제작에 혁신을 가져왔다. 그러나, 창작자의 동의 없이 수집된 데이터가 학습에 활용되거나, 원 저작물과 유사한 결과물이 무분별하게 활용되는 등 저작권 침해 가능성이 심화되고 있다. 이를 해결하기 위해 ViT 기반 합성 이미지 검사기와 저작권 판례 검색기를 구축하였고, LangGraph를 이용하여 LLM 에이전트가 이를 활용하여 검색 증강 생성을 수행하도록 구현하였다. 실험 결과 이미지 검사기를 통하여 사용자가 의심되는 모방 이미지에 대하여 정량적인 판단 기준을 제공받을 수 있으며, 검색 증강 생성을 통하여 사용자가 문맥적 일관성과 법률적 연관성이 높은 응답을 받을 수 있음을 확인하였다. 본 시스템의 AI 생성 여부 판별 및 저작권 침해 분석 에이전트 시스템을 통해 생성형 AI 시대의 지식재산권 보호 체계 구축에 기여하고자 한다.

I. 서 론

생성형 인공지능(Generative AI)의 발전은 음악, 미술 등 다양한 분야의 콘텐츠 제작 방식에 혁신을 가져왔다. 그러나 창작자의 동의 없이 수집된 데이터가 학습에 활용되거나, 원 저작물과 유사한 결과물이 무분별하게 활용되는 등 저작권 침해 가능성이 심화되고 있다[1,2].

본 연구는 생성형 AI 기반 이미지의 저작권 침해 여부를 기술적으로 판별하고, 법률적 자문을 제공하는 LangGraph 기반 LLM 에이전트를 제안한다. 제안하는 시스템을 통하여 AI 생성물 여부에 대한 판별을 거쳐 저작권 침해 분석을 통해 창작자의 권리 보호와 침해 대응을 지원함으로써, 생성형 AI 시대의 지식재산권 보호 체계 구축에 기여하고자 한다.

II. 관련 연구

합성 이미지 탐지. Utkarsh Ojha 등[3]은 CILP 모델에서 추출한 사전 학습된 Frozen ViT 특징 추출기와 단순 선형 분류기만을 활용하여, 다양한 생성 모델에 대해 우수한 일반화 성능을 보이는 합성 이미지 탐지 모델을 제안하였다.

RAG 기반 법률 자문 시스템. Song-Yeon Kim 등[4]은 입력된 사실 관계를 바탕으로 관련 법령과 주장·증명책임을 도출하는 모델을 제안하였다. 해당 모델은 공개된 판례 데이터에 법학 전공자가 주석을 추가하여 데이터셋을 구축하고, 이를 RAG 기법과 결합하여 답변을 생성한다.

차별성. 본 연구는 ViT 기반 이미지 탐지 결과를 LLM 에이전트의 응답 생성에 적용하고, 저작권 판례 검색을 통한 법률 자문을 통합 제공함으로써 AI 생성물 저작권

침해 의심 사례에 실질적 대응을 지원한다는 점에서 차별성을 가진다.

III. 시스템 설계

본 논문에서 제안하는 모델은 이미지 검사 모듈과 저작권 판례 검색 모듈로 구성되며, 이들은 LangGraph 기반의 에이전트 구조로 상호 연동된다. LangGraph의 명시적인 워크플로우 로직 제어를 통하여 시스템의 신뢰성 및 설명 가능성을 향상시켰으며, 그래프 구조의 워크플로우로 다양한 이미지 도메인 및 AI 생성 여부를 고려한 다양한 전략을 사용할 수 있도록 설계하였다.

이미지 검사 모듈. Swin Transformer 구조를 기반으로 입력 이미지의 AI 생성 여부를 판별한다. 저작권 침해 분석을 위해 Multimodal 접근법을 적용하여 세 가지 영역에서 유사도를 계산하여 저작권 침해 가능성을 정량적으로 산출한다: (1) ImageNet으로 사전 훈련된 ResNet50을 사용하여 2,048 차원 특징 벡터를 추출하고, 이를 바탕으로 코사인 유사도를 계산하여 의미적 유사성을 계산한다. (2) Wang 등[5]이 제안한 SSIM 알고리즘을 통해 휘도, 대비, 구조 정보를 종합한 시각적 유사성을 정량화 한다. (3) RGB 각 채널별 256-bin 히스토그램을 계산하고, Pearson 상관계수를 산출하여 색상 분포의 일치도를 평가한다.

저작권 판례 검색 모듈. 한국저작권위원회에서 발간한 저작권 상담사례집¹을 주요 데이터 소스로 수집하여 LLM 에이전트가 법률 자문 과정에서 판례 및 법률 조항을 활용하도록 하였다. 수집한 데이터를 GPT-4o-mini 모델을 활용하여 전처리하고 임베딩 벡터를 생성하였다.

¹<https://www.copyright.or.kr/business/law-case-counsel/index.do>

*아주대학교 소프트웨어학과 교수, 교신저자, kasohn@ajou.ac.kr

백터	2023 년 3 월, 172 회 서쿄에서 한 블루 아카이브 동인부스가 AI 그림을 트레이싱하고 굿즈를 만들어 논란이 되었다. 트레이싱이라고는 하나, 사람 손으로 새로이 그렸음에도 어색한 부분이 있어 걸렸다. 해당 굿즈들은 환불처리 진행중이다.
질문	제가 만든 블루아카이브 팬아트가 AI 트레이싱을 거쳐 굿즈로 판매되었습니다. 비록 일부 사람이 다시 그렸지만 여전히 제 원작을 기반으로 하고 있습니다. 이런 경우 저작권 침해를 주장하여 판매 중단과 손해배상을 청구할 수 있을까요?
참고 문서	Q: 21. 저작권 보호기간은 얼마나 되나요? A: 우리 저작권법은 특별한 규정이 있는 경우를 제외하고는 저작자가 생존하는 동안과 사망한 후 70년까지 저작권 보호기간을 인정 한다. 저작권은 저작물을 창작한 때부터 발생하지만 법에서 정한 보호기간이 경과하면 공유저작물(Public Domain)이 되어 누구나 자유로이 이용할 수 있다. 우리 저작권법은 특별한 규정이 있는 경우를 제외하고는 저작자가 생존하는 동안과 사망한 후 70년까지 저작권 보호기간을 인정하고 있으며, 보호기간은 저작자가 사망한 다음 해부터 기산한다.
생성 결과	귀하의 블루아카이브 팬아트가 AI 트레이싱을 거쳐 굿즈로 판매된 경우, 저작권 침해를 주장할 수 있는 가능성이 있습니다. 저작권법에 따르면, 원작자의 저작물에 대한 권리는 저작자가 생존하는 동안과 사망 후 70년까지 보호됩니다. 귀하의 팬아트가 원작을 기반으로 하여 창작된 경우, 해당 팬아트의 저작권은 귀하에게 있으며, 이를 무단으로 사용한 경우 저작권 침해가 성립할 수 있습니다.

표 1. 저작권 침해 관련 사례 및 질의응답 생성 결과 예시

생성된 백터는 FAISS (Facebook AI Similarity Search) 기반 백터 데이터베이스(Vector DB)에 저장하여 고속 유사도 검색이 가능하도록 하였다. 검색 결과는 LLM 에 전달되어 법률 연관성을 고려한 응답을 생성한다.

IV. 실험 결과

프로토타이핑. 설계한 에이전트 시스템을 검증하기 위하여 프로토타입 모델을 개발하여 실험 및 검증을 진행하였다. 프로토타입 모델은 gpt-4o-mini 생성 모델을 기반으로 하며, streamlit GUI 를 이용하여 그림 1 과 같이 사용자 인터페이스를 구현하였다.



그림 1. 프로토타입 모델 인터페이스

저작권 상담 사례 검색. 저작권 침해와 관련된 10 개의 국내 및 해외 사례를 수집하였고, 사건 내용을 토대로 GPT-5 모델을 통해 법률 질문을 생성하였다. 표 1 과 같이, LLM 에이전트는 검색 문서를 적절하게 활용하여 최종 응답 생성에 활용하는 볼 수 있다.

이미지	실제 이미지	AI 모방 이미지	같은 도메인의 다른 실제 이미지	다른 도메인의 실제 이미지
의미적 유사성	-	0.8444	0.5232	0.6112
시각적 유사성	-	0.2082	0.1837	0.1348
색상분포 일치도	-	0.7023	0.4211	0.2884

표 2. 이미지 유사도 분석 결과

이미지 유사도 검사. 명화 및 애니메이션의 실제 이미지와 모방 이미지의 쌍으로 이루어진 총 20 장의 이미지를 준비하여 실제 이미지와 모방 이미지 간 이미지 유사도 검사 결과를 확인하였다. 표 2 를 보면,

실제 이미지를 모방한 이미지의 유사성이 실제 이미지와 관련 없는 다른 이미지보다 높았음을 확인하였다. 따라서 이미지 검사 모듈이 유사한 이미지를 판단할 수 있으며, 이를 통해 창작자가 의심되는 이미지에 대하여 정량적인 판단 기준을 제공받을 수 있음을 의미한다.

V. 결론 및 향후 연구

본 연구에서는 생성형 AI 기반 이미지의 저작권 침해 여부를 기술적으로 판별하고, 법률적 자문을 제공하는 LangGraph 기반 LLM 에이전트 시스템을 제안하였다. 이미지 검사 모듈과 저작권 판례 검색 모듈을 통하여 정량적 판단 기준과 함께 문맥적 일관성과 법률적 연관성이 높은 답변을 얻을 수 있음을 확인하였다. 후속 연구로 캐릭터, 회화, 소설 등과 같은 다양한 유형의 콘텐츠 도메인별 침해 유형을 분류하고 대응 전략을 추천하는 알고리즘을 고도화 할 수 있을 것이다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 인공지능융합혁신인재양성사업 연구 결과로 수행되었음(IITP-2025-RS-2023-00255968)

참 고 문 헌

- [1] 황지은. (2025). LLM 기반 생성 콘텐츠의 저작권 및 표절 관련 이슈 연구. 교양학연구, 12(1), 189-209. 10.55450/hilas.2025.12.1.189
- [2] Li, O., Cai, J., Hao, Y., Jiang, X., Hu, Y., & Feng, F. (2025, July). Improving synthetic image detection towards generalization: An image transformation perspective. In Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1 (pp. 2405-2414).
- [3] Ojha, U., Li, Y., & Lee, Y. J. (2023). Towards universal fake image detectors that generalize across generative models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 24480-24489).
- [4] 김송연, 김창현, 이상대, 박혜진, 김영민. (2024-12-18). 주장증명책임을 안내하는 RAG 기반 법률 LLM. 한국정보과학회 학술발표논문집, 전남.
- [5] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, 13(4), 600-612. doi:10.1109/TIP.2003.81986