

저작권 이슈 브리프



COPYRIGHT ISSUE BRIEF

Weekly Report
2026. 3-4



한국저작권위원회
KOREA COPYRIGHT COMMISSION

본 보고서는 EC21R&C(컨설팅사)에서 작성하였고, 국내외 저작권 기술·산업 동향을 조사한 자료로 한국저작권위원회 의견이 반영되어 있지 않습니다.



SUMMARY

산업/기업

기술

산업 넷플릭스와 아마존, 영화 제작에 AI 도입 추진

▶ 넷플릭스가 2026년 3월 미국의 AI 스타트업 인터포지티브를 약 6억 달러에 인수하며 영화 제작 자동화 기술을 확보했다. 인터포지티브는 연속성 오류 수정, 조명 보정, 배경 수정 등 영화 제작 후반작업 단계에 특화된 AI 모델을 개발한 기업으로, 해당 모델은 허가 없이 데이터를 학습하지 않아 권리 침해 가능성을 최소화한다. 아마존 역시 색 보정, 사운드 믹싱, 시각 효과 등 후반작업 과정에서의 AI 활용을 본격화하고 있는데, 넷플릭스처럼 전문 기업을 인수하는 방식이 아닌, 자사의 기술과 역량을 활용하는 전략을 추진 중이다. 이처럼 현재 영화 제작 산업에서는 AI를 통해 제작 비용을 절감하고 제작 기간을 단축하는 움직임이 본격화되는 모습이다. 그러나 일각에서는 일부 소비자들이 여전히 AI 생성 콘텐츠에 대한 거부감을 갖고 있기 때문에 실제 도입은 예상보다 늦어질 수 있다는 전망도 제기되고 있다.

산업 애플 뮤직, 공급자가 직접 AI 사용 여부를 입력하는 'AI 투명성 태그' 도입

▶ 생성형 AI 음악 도구 확산으로 AI 생성 음원이 스트리밍 플랫폼에 빠르게 유입되고 있다. 글로벌 음악 스트리밍 플랫폼 디저에 따르면 AI 생성 음원의 하루 유입량은 2025년 1월 약 1만 건에서 2026년 1월 6만 건 이상으로 증가하였고, 유입된 AI 생성 음원 중 약 85%가 스트리밍 조작(abusing) 행위 등 악용 사례도 늘고 있다. 이에 애플 뮤직은 2026년 3월, 음원 등록 시에 AI 활용 여부를 표시하는 '투명성 태그' 체계를 도입하여, 아트워크, 트랙, 작곡, 뮤직비디오 등 4개 영역에 대해 공급자가 직접 AI 사용 여부를 입력하는 방식으로 운영을 변경하였다. 디저는 자체 기술을 활용하여 AI 생성 음원을 탐지하고, 스포티파이는 DDEX 표준을 도입하는 등 대응 방식은 플랫폼별로 각기 다른 양상을 보이고 있다. 향후에는 업계 공통의 분류 기준을 마련하고 이를 권리 관리 체계와 연계하는 것이 주요 과제가 될 것으로 보인다.

산업 유튜브, 유사성 탐지 기능의 적용 범위를 공적 인물까지 확대

▶ 유튜브가 2025년 4월 딥페이크에 대응하기 위해 유사성 탐지 기능을 시범 도입한 데 이어, 2026년 3월에는 정부 관계자·언론인·정치 후보자 등 공적 인물까지 적용 범위를 확대했다. 이 기능은 등록자의 본인 확인을 거쳐 생성된 유사성 템플릿을 기준으로, 다른 사용자가 새로 업로드한 영상에서 유사한 시각적 특징을 탐지하는 방식이다. 탐지된 영상은 자동으로 삭제되지 않으며, 등록자가 직접 검토한 뒤 개인정보 보호 정책을 위반한다고 판단될 경우 삭제를 요청하는 구조이다. 유튜브는 향후 음성의 유사성까지 탐지할 수 있도록 기능을 확장할 계획이다.



저작권 이슈 브리프

SUMMARY

산업/기업

기술

산업 캔바의 매직 레이어스 출시와 AI 생성 이미지 편집 환경의 변화

▶ 캔바가 하나의 이미지 안에 포함된 배경·그래픽·텍스트 등의 구성 요소를 자동으로 인식하고, 각 요소를 독립적으로 편집할 수 있는 레이어로 분리해 주는 기능인 '매직 레이어스'를 출시했다. 매직 레이어스 활용 시, 구성 요소가 서로 겹쳐 가려진 영역은 AI가 주변 이미지 정보를 바탕으로 자동으로 채워 복원하며, 이미지에 포함된 텍스트는 편집이 가능한 문자 데이터로 변환된다. 이러한 기능을 AI 생성 이미지에 적용하면, 프롬프트를 수정해 이미지 전체를 다시 생성하지 않고도 필요한 부분만 편집할 수 있어 작업 효율이 높아질 것으로 예상된다. 다만 기존에는 단일 레이어로 출력되던 AI 생성 이미지도 다층 레이어 구조로 변환할 수 있게 되면서, 레이어 구조만으로 AI 생성 여부를 판별하는 것이 어려울 수도 있다는 우려도 제기되고 있다.

산업 중국 AI 만화극 시장 확대와 저작권 관리 체계 변화

▶ 생성형 AI가 대본 설계부터 캐릭터 생성, 음성 합성에 이르는 콘텐츠 제작 전 과정에 개입되면서, 기존의 인간 창작자 중심의 저작권 체계와 저작물 심사 방식이 그대로 운영되기 어려워졌다. 이러한 한계를 보완하기 위해 최근 중국은, ▲ AI 생성 콘텐츠의 저작물성 판단 기준 확립, ▲ 학습 데이터와 소재 출처 검증, ▲ 플랫폼·창작자·도구 제공자 간 책임 배분, ▲ 콘텐츠 분류·표시 의무화를 통해 관리 체계를 개편하고 있다. 중국에서 이러한 변화는 특히 AI 만화극 산업에서 두드러지고 있지만, 향후 미디어 콘텐츠 산업 전반의 권리·책임 구조에 대한 재검토 논의로 이어질 것으로 보인다.

기술 주간 기술 동향

▶ 생성형 AI 활용 확산으로 저작권이 있는 이미지의 무단 사용이 급증하고 있음에도 여전히 창작자들이 자신의 저작물에 대한 저작권 침해를 입증하는 것은 쉽지 않다. 일례로, 2024년 사진작가 조 지론과 라인 파이낸셜 간 소송에서 지론은 자신의 저작물에 대한 침해 사실이 확인되었음에도 불구하고 라이선스 기록이 부재하다는 이유로 손해배상을 받지 못했다. 기존에는 이러한 문제를 해결하기 위한 기술적 방법으로 이미지 워터마킹 기술이 활용되어 왔으나, 이러한 기술은 간단한 편집만으로도 쉽게 제거될 수 있다는 점에서 한계가 분명히 존재한다. 이와 관련하여 최근 개발된 'SLICE' 기술은 기존 기술의 한계점을 일부 보완할 것으로 보인다. SLICE 기술은 이미지를 주제, 환경, 행동, 세부사항 4가지 요소로 분해하고 각각을 독립적인 공간 파티션에 고정하여, 변조된 부분을 정확히 탐지한다.



넷플릭스와 아마존, 영화 제작에 AI 도입 추진

메이저 미디어 기업들, AI를 통한 영화 제작 효율화 추진

• 넷플릭스, 영상 제작 AI 스타트업 인수

- 넷플릭스는 2026년 3월 미국의 배우 겸 영화 감독 벤 애플렉(Ben Affleck)이 공동 창업한 AI 스타트업 인터포지티브(InterPositive)를 약 6억 달러(원화 약 8,600억 원)¹⁾에 인수함
- 벤 애플렉은 'AI가 영화 제작 과정에서 노동력이 많이 소요되는 부분을 자동화함으로써 제작 비용을 낮추고 독립 영화 제작자의 진입 장벽을 낮출 것'²⁾이라고 언급한 바 있음
- 넷플릭스는 이미 자사의 오리지널 시리즈 '이터노트(The Eternaut)'에서 건물 붕괴 장면을 생성형 AI로 제작한 바 있음
- 인터포지티브는 기존 촬영분에서 ▲한 시퀀스 안에서의 연속성 검토 ▲누락된 장면 보완 ▲조명의 강도 및 각도 보정 ▲불필요한 물체(세트장 구조물, 붐 마이크, 배우가 아닌 스태프, 와이어 등) 제거 ▲배경 수정 등 후반작업에 특화된 AI 모델을 개발함
- 인터포지티브의 AI는 원작자의 허가 없이 영화 데이터를 학습하지 않으며, 새로운 콘텐츠를 자동 생성하지 않는다는 점에서 저작권 침해 위험을 최소화하는 구조를 갖추고 있음. 이는 배우나 세트장 등 영화의 기존 요소를 모두 대체하려고 시도하던 기존 생성형 AI와 차별화된 형태임.
- 인터포지티브의 AI는 감독이나 영화 제작자가 촬영 당일의 영상을 업로드하여 특정 프로젝트에 맞게 모델을 최적화할 수 있는 기능을 제공함. 이를 통해 촬영을 모두 마친 후 후반작업을 시작하던 기존 제작 방식보다 제작 기간을 대폭 줄일 수 있을 것으로 기대됨

• 넷플릭스의 조직 통합과 AI 거버넌스 체계 구축

- 이번 인수를 통해 벤 애플렉을 포함한 인터포지티브의 전 직원이 넷플릭스로 이동했으며, 이는 단순한 기술 이전을 넘어 AI 전문 인력과 노하우를 조직 내부로 완전히 통합하는 방식임
- 넷플릭스는 인터포지티브 인수 발표와 함께 기술, 콘텐츠 분야의 임원 토론 영상을 공개하여 AI가 창작 과정에서 어떤 역할을 할 것인지에 대한 내부 논의를 공유함
- 이러한 조직적 통합과 최고 경영진 차원의 공개 논의는 넷플릭스가 AI 기술을 단순한 영화 제작 도구가 아닌 자사의 콘텐츠 제작 철학과 조직 문화에 통합하는 것을 목표로 하고 있음을 시사함

• 아마존의 자체 AI 개발 및 AWS 통합 전략

- 한편, 아마존 MGM 스튜디오(Amazon MGM Studios, 이하 MGM)는 2026년 3월부터 색 보정, 사운드 믹싱, 시각 효과 등 후반작업의 전 과정을 자동화하는 자체 개발 AI 도구를 테스트하고 있으며, 올해 5월까지 초기 결과를 공개할 예정임

1) 1달러=1,435.40원(2026.03.03, KEB 하나은행 매매기준율 적용), 이하 동일 기준 적용

2) Alex Weprin, "Ben Affleck Quietly Founded a Filmmaker-Focused AI Tech Company. Netflix Just Bought It.", The Hollywood Reporter, 2026.03.05., <https://www.hollywoodreporter.com/business/digital/ben-affleck-ai-netflix-1236521806/>

- 아마존의 AI 도구는 자사 AWS(Amazon Web Services) 인프라를 기반으로 엔트로픽, 오픈AI 등 다중 LLM 제공자를 통합하는 구조로 설계됨
- 이는 단일 모델에 의존할 경우 발생할 수 있는 기술적·전략적 리스크를 최소화하기 위한 전략으로 분석됨
- 아울러, 아마존의 자체 클라우드 컴퓨팅 역량과 AI 연구 조직을 활용하여 콘텐츠 제작부터 배포까지 수직 통합을 강화하려는 전략으로도 볼 수 있음

AI 기반 제작비 절감 전략과 주요 도전과제

• 후반작업 자동화를 통한 비용 절감 전략

- 넷플릭스와 아마존은 AI 도구를 활용하여 영화를 제작할 경우, 콘텐츠가 관객에게 도달하기 전까지 소요되는 상당한 예산을 절감할 수 있다고 강조함
- 넷플릭스는 인터포지티브의 AI 도구가 배우의 연기 자체가 아닌 영화 제작 기법에 중점을 두고 있다고 밝혔으며, AI를 활용해 반복적인 편집 과정을 간소화하고 제작 소요 기간을 줄이는 것을 목표로 한다고 부연함
- 한편, 아마존은 사전제작 단계의 시퀀스 시각화, 제작 단계(카메라 작업 및 특수효과 프리뷰), 후반작업 단계(색 보정·사운드 믹싱·시각 효과 개선) 등 제작 파이프라인 전반에 AI를 통합하고 있음

• AI 콘텐츠에 대한 관객 수용성과 품질 우려

- AI 생성 콘텐츠에 대한 관객 수용성은 여전히 중요한 변수로 남아 있는데, 특히 관객이 품질 문제를 이유로 AI를 활용해 만든 최종 결과물에 대해 부정적 평가를 할 가능성도 존재함
- 또한, AI 도구가 제작 효율성을 높이더라도 감독이나 제작자의 의도를 파악하고 영상을 보정하는 것은 아닐 수 있기 때문에, 최종 콘텐츠의 창의적 표현이나 감정적 깊이가 저하될 수 있음.
- 업계 관계자들은 이러한 불확실성을 고려할 때, 실제 제작 환경에서의 AI 도입이 예상보다 느리게 진행될 수 있다고 설명함

참고문헌

- Alex Weprin, “Ben Affleck Quietly Founded a Filmmaker-Focused AI Tech Company. Netflix Just Bought It.”, The Hollywood Reporter, 2026.03.05., <https://www.hollywoodreporter.com/business/digital/ben-affleck-ai-netflix-1236521806/>
- Prashant Jagtap, “Amazon Launches AI Production Tools for Film and TV: Hollywood Faces Job Disruption Concerns”, 2026.02.05., Quill Circuit, <https://www.quillcircuit.com/blog/amazon-launches-ai-production-tools-for-film-and-tv-hollywood-faces-job-disruption-concerns>



애플 뮤직, 공급자가 직접 AI 사용 여부를 입력하는 'AI 투명성 태그' 도입

음원 플랫폼 내 AI 생성 음원 유입 확대

• AI 생성 음원 유입 확대 및 악용 사례 증가

- 생성형 AI 음악 도구가 확산됨에 따라 AI가 생성한 음원이 스트리밍 플랫폼에 빠르게 유입되고 있음
- 글로벌 음악 스트리밍 플랫폼인 디저(Deezer)는, 자사 플랫폼에 유입되는 AI 생성 음원이 2025년 1월 하루 약 1만 건에서 2026년 1월 기준 6만 건 이상으로 증가하였다고 발표함³⁾
- AI 생성 음원 재생 건수 중 약 85%는 저작권료를 부당하게 수취하기 위해 재생 횟수를 인위적으로 부풀리는 것으로 확인됨¹⁾

• AI 음원에 대한 투명성 확보 과제 대두

- 현재 음원을 스트리밍 플랫폼에 등록할 때 공급자가 입력하는 정보는 장르, 크레딧*, 발매일 등 일반적인 정보에 한정되어 있어, AI가 활용되었는지를 별도로 표시하거나 확인할 수 있는 항목은 없음
- 음원의 AI 활용 여부가 공개되지 않을 경우, 해당 음원의 저작권이 누구에게 귀속되는지 판단하기 어렵고 저작권료 정산 과정에서도 분쟁이 발생할 가능성이 높아짐
- 이러한 가운데 최근, 주요 스트리밍 플랫폼을 중심으로 AI 생성 음원을 유통 단계에서 식별하고 공개하려는 움직임이 나타나고 있으며, 그 방식은 플랫폼마다 다르게 전개되고 있음
- 일례로 애플 뮤직(Apple Music)은 2026년 3월, 음원 등록 단계에서 AI 활용 여부를 표시하는 '투명성 태그(Transparency Tags)**'를 도입함

* 크레딧(credits): 음원의 창작 및 제작에 참여한 인물과 역할을 기록한 정보로, 작곡가, 작사가, 가수, 연주자, 프로듀서, 엔지니어 등이 포함됨

** 투명성 태그(Transparency Tags): 애플 뮤직이 2026년 3월 도입한 AI 콘텐츠 공개 체계로, 음반사와 유통사가 콘텐츠 납품 시 AI 활용 여부를 태그 형태로 표시하는 자기신고 방식의 표기 체계

3) Murray Stassen, "Apple Music launches AI transparency tags — but only if labels and distributors declare them", Music Business Worldwide, 2026.03.04., <https://www.musicbusinessworldwide.com/apple-music-launches-ai-transparency-tags-but-only-if-labels-and-distributors-choose-to-declare-them/>

애플 뮤직이 도입한 투명성 태그 정책과 타 플랫폼의 접근 방식 비교

• 투명성 태그, 4대 범주에 대한 AI 활용 여부

- 애플 뮤직의 투명성 태그 정책은 AI가 활용된 영역을 아트워크(artwork), 트랙(track), 작곡(composition), 뮤직비디오(music video)의 4개 범주로 나누어, 각 영역에 AI가 관여했는지 여부를 표기하도록 하고 있음

[표1] 애플 뮤직이 도입한 투명성 태그의 분류 체계

태그 유형	적용 단위	적용 기준
아트워크(artwork)	앨범 단위	AI가 정지 또는 모션 그래픽 아트워크의 상당 부분을 생성한 경우
트랙(track)	트랙 단위	AI가 음원 녹음의 상당 부분을 생성한 경우
작곡(composition)	트랙 단위	AI가 가사 또는 기타 작곡 요소의 상당 부분을 생성한 경우
뮤직비디오(music video)	앨범 또는 단독	AI가 영상 요소의 상당 부분을 생성한 경우

출처: Stuart Dredge, "Apple Music adds 'transparency tags' for uses of AI with music", Music Ally, 2026.03.05., <https://musically.com/2026/03/05/apple-music-adds-transparency-tags-for-uses-of-ai-with-music/>

- 각 태그는 AI가 해당 요소의 '상당 부분(material portion)'에 활용된 경우 표기하도록 되어 있으나, 어느 정도를 '상당 부분'으로 볼 것인지에 대한 판단은 음반사와 유통사에 위임됨
- 투명성 태그는 현재 선택 사항이나, 향후 신규 음원 등록 시 필수 적용으로 전환될 예정임. 태그가 입력되지 않은 경우에는 AI가 활용되지 않은 콘텐츠로 간주됨
- 기존의 일반적인 음원 정보와 마찬가지로, AI 활용 여부 역시 음반사와 유통사가 직접 판단하여 입력하는 방식으로 운영할 수 있다는 설명임
- 애플 뮤직은 이번 투명성 태그 정책을 통해 음원 시장 내 AI 사용과 관련된 정보를 축적하고, 이를 향후 신규 정책 수립에 활용할 수 있을 것이라고 설명함

• 여타 스트리밍 플랫폼의 AI 음원 대응 방식

- 앞서 언급한 디저는 자체 기술로 AI 생성 음원을 탐지하고 있으며, 2026년 3월 기준 누적 1,340만 건 이상의 AI 생성 음원을 식별하였다고 밝힘⁴⁾
- 특히 수노(Suno), 유디오(Udio) 등 주요 생성형 AI 음악 도구를 사용한 음원은 대부분 식별 가능하다고 강조하였으며, 해당 기술을 외부 저작권 관리 기관에도 제공하는 단계로 확장 중임
- 한편 세계 최대 음악 스트리밍 플랫폼인 스포티파이(Spotify)는 2025년 9월 디지털 음원 유통 국제 표준화 기구(Digital Data Exchange, DDEX)*의 AI 기여 표시 기준을 도입함
- DDEX 기준은 보컬, 악기 연주, 후반작업 등 제작 단계별 AI 활용 여부를 구분하고, 이를 표준화된 메타데이터 형식으로 음원 크레딧에 포함함으로써 공통된 방식의 AI 표시 체계를 구축하는 것을 목표로 함

* 디지털 데이터 표준화(Digital Data Exchange, DDEX): 음악 산업의 디지털 공급망에서 사용되는 메타데이터 및 거래 표준을 개발하는 국제 기구

4) Murray Stassen, "Apple Music launches AI transparency tags — but only if labels and distributors declare them", Music Business Worldwide, 2026.03.04., <https://www.musicbusinessworldwide.com/apple-music-launches-ai-transparency-tags-but-only-if-labels-and-distributors-choose-to-declare-them/>

투명성 태그의 실효성 여부와 향후 도전 과제

• 자기신고에 기반한 정보 공개 방식의 한계

- 애플 뮤직의 투명성 태그는 공급자가 태그를 제대로 입력했는지 확인하거나 강제할 수 없어, 태그가 없을 경우 AI 활용 여부가 드러나지 않는 한계가 있음
- 또한 AI 활용에 해당하는지의 판단도 공급자에 전적으로 맡기고 있어, AI를 활용했더라도 공급자에 따라 태그 부착 여부가 달라질 수 있는 등 정보 공개의 일관성을 확보하기 어려움
- 자기신고 방식에 의존하는 투명성 태그가 실효성을 갖추려면 미신고나 허위신고에 대한 제재 기준과, 신고 내용을 교차 확인할 수 있는 검증 절차가 뒷받침되어야 함
- 애플 뮤직은 부정 재생에 대해 이미 금전적 제재를 부과하고 있어, AI 활용 여부를 미신고한 경우에도 유사한 제재가 도입될 가능성이 있음

• AI 생성 음원에 대한 공통 표기 기준의 필요성

- 앞서 살펴본 바와 같이 스트리밍 플랫폼마다 음원의 AI 활용 여부를 공개하는 방식이 다른 만큼, 같은 음원이라도 플랫폼에 따라 AI 판정 결과가 달라질 수 있음
- 이로 인해, 음원 스트리밍 업계 내 AI 활용 여부에 대한 통일된 기준을 마련하는 것이 주요 해결 과제가 됨. 향후 투명성 체계와 권리 관리를 연계하는 방안에 대한 논의도 확대될 수 있을 것으로 전망됨

참고문헌

- Murray Stassen, “Apple Music launches AI transparency tags — but only if labels and distributors declare them”, Music Business Worldwide, 2026.03.04., <https://www.musicbusinessworldwide.com/apple-music-launches-ai-transparency-tags-but-only-if-labels-and-distributors-choose-to-declare-them/>
- Stuart Dredge, “Apple Music adds ‘transparency tags’ for uses of AI with music”, 2026.03.05., Music Ally, <https://musically.com/2026/03/05/apple-music-adds-transparency-tags-for-uses-of-ai-with-music/>
- Crystal Koe, “Apple Music’s new ‘Transparency Tags’ aim to flag AI-generated content – but labels have to self-report”, Music Tech, 2026.03.07., <https://musictech.com/news/industry/apple-music-transparency-tags-ai/>



저작권 이슈 브리프

SUMMARY

산업/기업

기술

유튜브, 유사성 탐지 기능의 적용 범위를 공적 인물까지 확대

유튜브, 딥페이크 대응 위해 유사성 탐지 기능의 적용 범위 확대

• 유튜브, 공적 인물의 초상을 보호할 도구가 필요하다고 판단

- 유튜브(YouTube)는 2025년 4월, 딥페이크에 대응하기 위해 유사성 탐지(Likeness Detection)* 기능을 시범 도입하고, 같은 해 10월 유튜브 파트너 프로그램(YPP)** 소속 일부 크리에이터를 대상으로 순차 출시를 시작함
- 이후 2026년 3월 10일에는 정부 관계자·언론인·정치 후보자를 대상으로 시범 운영을 시작하여 적용 범위를 공적 인물까지 확대하였음
- 유튜브는 뉴스 보도와 공적 토론의 중심에 있는 인물일수록 이들의 초상을 보호할 수 있는 도구가 필요하다는 취지에서 적용 범위를 확대했다고 설명함¹⁾
- 이번 확대는 유사성 탐지 기능이 공적 인물에게 적합한지 점검하기 위한 시범 운영으로, 유튜브는 향후 수개월 내에 해당 기능의 접근 범위를 확대할 계획도 밝힘

* 유사성 탐지(Likeness Detection): 등록자의 얼굴 정보를 바탕으로 유튜브 영상에서 등록자와 유사한 시각적 특징을 탐지하는 기능

** 유튜브 파트너 프로그램(YouTube Partner Program, YPP): 일정 요건을 충족한 크리에이터가 광고 수익 배분, 채널 기능, 권리 보호 도구 등에 접근할 수 있도록 한 프로그램

유튜브 유사성 탐지 기능의 작동 구조

• 유사성 탐지 기능의 등록 절차 및 데이터 처리 구조

- 유튜브의 유사성 탐지 기능은 기존 콘텐츠 ID(Content ID) 시스템과 동일한 기술 원리를 기반으로 하며, 탐지 대상이 저작물 대신 인물의 초상이라는 점에서 차이가 있음
- 유사성 탐지 기능을 이용하려면 등록자는 정부가 발급한 신분증과 얼굴을 촬영한 짧은 영상을 제출하여 본인 확인을 거쳐야 함. 이는 타인의 얼굴 정보가 무단으로 등록되는 것을 방지하기 위한 절차임
- 유튜브는 제출된 얼굴 영상과 등록자가 본인의 채널에 업로드한 콘텐츠 속 얼굴 이미지를 바탕으로 유사성 템플릿(Likeness Template)*을 생성함. 이 템플릿은 이후 다른 사용자가 업로드한 영상을 탐지할 때 비교 기준으로 활용됨
- 등록자의 유사성 템플릿에는 고유 식별자가 부여되며, 마지막으로 로그인한 시점부터 최대 3년간 보관된 뒤 삭제됨. 단, 등록자가 동의를 철회하거나 계정을 삭제한 경우에는 즉시 삭제됨
- 유사성 탐지를 설정하기 위해 제출된 데이터는, 등록자가 동의하지 않는 한 구글의 생성형 AI 모델을 학습시키는 데 사용되지 않음

* 유사성 템플릿(Likeness Template): 등록자가 제출한 얼굴 영상과 본인 채널의 콘텐츠를 바탕으로 생성되는 얼굴 비교 기준으로, 이후 다른 영상에서 등록자의 얼굴이 사용되었는지를 판별하는 데 활용됨

1) Colin Kirkland, "YouTube Brings AI Likeness Detection To Journalists, Civic Leaders", MediaPost, 2026.03.10., <https://www.mediapost.com/publications/article/413389/youtube-brings-ai-likeness-detection-to-journalist.html>

• 등록자의 유사성 템플릿을 기반으로 탐지하는 방식

- 유사성 탐지 기능은 등록된 유사성 템플릿을 기준으로, 새로 업로드되는 영상을 각 영상당 1회씩 스캔하는 방식으로 작동함
- 탐지 과정에서 비등록자의 얼굴도 함께 스캔되나, 등록자의 유사성 템플릿과 일치하지 않는 데이터는 즉시 삭제됨
- 스캔 결과 잠재적으로 일치하는 영상이 발견되면, 등록자가 직접 검토할 수 있도록 해당 영상 목록을 제공함

• 등록자와 유튜브가 검토한 뒤 삭제하는 절차

- 잠재적으로 일치하는 영상이 발견되더라도 자동 삭제로 이어지지는 않음. 등록자가 영상을 검토한 뒤, 유튜브의 개인정보 보호 정책을 위반했다고 판단하면 유사성 삭제 요청을 제출할 수 있음
- 다른 채널이 원본 영상을 무단으로 재사용한 경우에는 유사성 삭제가 아닌 저작권 침해 삭제 요청 절차를 통해 처리됨
- 또한 유튜브는 유사성 삭제 요청이 접수된 경우에도 해당 영상이 패러디·풍자에 해당하는지를 함께 검토하는데, 이는 표현의 자유를 고려하기 위한 것이라고 밝힘²⁾

유사성 탐지 기능의 향후 전망과 과제

• 기술적 수단의 고도화와 법적 체계 확립의 병행 필요성

- 현재 유사성 탐지 기능은 얼굴의 시각적 유사성을 찾는 데 한정되어 있어, 음성을 합성하여 만든 AI 산출물에는 대응하기 어려움. 이에 유튜브는 2026년 내로 음성 유사성 탐지까지 기능을 확장할 계획임
- 또한 유튜브는 초상권을 보호하는 데 기술적 수단만으로는 충분하지 않다는 입장을 밝히고 있음. 이와 함께 연방 차원에서 퍼블리시티권을 확립하는 딥페이크 방지법(NO FAKES Act)과 같은 법적 체계를 마련해야 한다는 입장임²⁾
- 이는 플랫폼 차원의 탐지 기능이 초상을 보호하는 하나의 수단이 될 수 있으나, 권리 보호의 실효성을 확보하기 위해서는 법적·제도적 뒷받침이 함께 갖추어져야 한다는 점을 시사함

* 딥페이크 방지법(NO FAKES Act): 개인의 얼굴 음성 등 초상을 시로 무단 재현하는 행위를 규제하기 위해 미국 연방 차원에서 퍼블리시티권을 확립하려는 법안

참고문헌

- Colin Kirkland, "YouTube Brings AI Likeness Detection To Journalists, Civic Leaders", MediaPost, 2026.03.10., <https://www.mediapost.com/publications/article/413389/youtube-brings-ai-likeness-detection-to-journalist.html>
- Youtube Help, "Likeness detection on YouTube", 2026.03.19. 접속 기준, <https://support.google.com/youtube/answer/16440338?hl=en>
- Amjad Hanif, "Expanding likeness detection to civic leaders and journalists", Youtube Official Blog, 2026.03.10., <https://blog.youtube/news-and-events/expanding-likeness-detection-civic-leaders-journalists/>

²⁾ Amjad Hanif, "Expanding likeness detection to civic leaders and journalists", Youtube Official Blog, 2026.03.10., <https://blog.youtube/news-and-events/expanding-likeness-detection-civic-leaders-journalists/>



저작권 이슈 브리프

SUMMARY

산업/기업

기술

캔바의 매직 레이어스 출시와 AI 생성 이미지 편집 환경의 변화

캔바, 이미지 레이어 자동 분리·편집 기능인 매직 레이어스 출시

- 매직 레이어스, 이미지 구성 요소를 자동으로 인식하여 레이어로 분리·편집 지원
 - 클라우드 기반 디자인 플랫폼인 캔바(Canva)가 매직 레이어스(Magic Layers)를 출시함. 매직 레이어스는 AI가 이미지를 분석하여 구성 요소를 자동으로 인식하고, 각 요소를 독립된 레이어로 분리하여 편집할 수 있게 해주는 기능임
 - 일반적으로 이러한 구성 요소들이 포함된 이미지가 파일(JPEG, PNG 등)로 저장되면, 구성 요소들은 하나로 병합되어 개별적으로 편집할 수 없는 상태가 됨
 - 하나의 이미지는 여러 장의 투명 필름을 겹쳐 놓은 것처럼 레이어가 겹겹이 쌓여 구성됨. 사용자는 원하는 레이어만 선택하여 해당 구성 요소를 독립적으로 편집할 수 있음
 - 이미지의 구성 요소들이 하나로 합쳐진 상태를 단일 레이어 이미지라 하고, 구성 요소가 각각 별도의 레이어로 분리된 상태를 다층 레이어 구조라 함. 매직 레이어스는 단일 레이어 이미지를 분석하여 다층 레이어 구조로 변환해 주는 기능임
 - 현재 캔바는 유료 구독자를 대상으로 매직 레이어스를 미국·영국·캐나다·호주에서 시험 운영 중임
 - * 구성 요소: 배경, 텍스트, 그래픽, 인물 등 하나의 이미지 안에 포함된 개별 시각적 요소
 - ** 레이어: 구성 요소를 독립적으로 편집할 수 있도록 분리해 놓은 단위를 뜻하며, 하나의 레이어에는 하나의 구성 요소가 담김

[그림1] 눈에 보이는 이미지(좌), 실제 레이어 구조(우)



출처: 캔바 공식 홈페이지 화면 캡처, 2026.03.19. 접속 기준, <https://www.canva.com/magic-layers/>

매직 레이어스의 작동 방식과 주요 기능

• 구성 요소 자동 구분, 중첩 영역 보정, 텍스트 변환의 세 가지 기능으로 작동

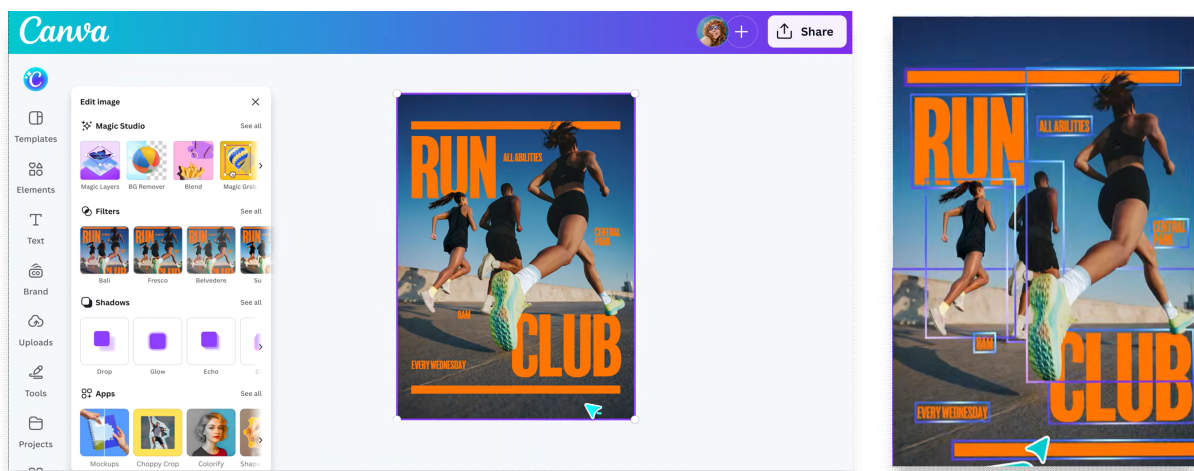
- 사용자는 편집하려는 이미지를 캔바의 편집 화면에 업로드한 뒤, 매직 레이어스를 실행할 수 있음
- 매직 레이어스는 사용자가 업로드한 이미지를 분석하여 배경·그래픽·텍스트 등의 구성 요소를 각각 구분함. 구분이 완료되면 각 요소는 독립된 레이어로 분리되며, 사용자는 각 레이어를 제거하거나 색상·크기를 변경하는 등 개별적으로 편집할 수 있음
- 구성 요소가 서로 겹쳐서 가려져 있던 영역은 AI가 주변 이미지 정보를 바탕으로 자동으로 채워 복원함
- 예를 들어, 배경 앞에 인물이 겹쳐져 있던 이미지에서 인물만 별도의 레이어를 분리하면, 인물이 있던 자리는 공백으로 남게 됨. 매직 레이어스는 이 공백을 주변 배경과 유사한 색상·질감으로 자동으로 채움
- 단일 레이어 이미지에 포함된 텍스트의 경우, 겉으로는 글자처럼 보이지만 실제로는 이미지의 일부로 처리되어 내용을 개별적으로 편집하기 어려웠음
- 매직 레이어스는 이 텍스트를 별도의 레이어로 분리하는 과정에서 편집 가능한 문자 데이터로 변환하며, 폰트·크기·색상 등의 세부 속성까지 직접 수정할 수 있는 상태가 됨
- 다만, 매직 레이어스는 그래픽 디자인이나 일러스트레이션 기반의 이미지에서 가장 잘 작동하며, 사실적 사진에서는 구성 요소 분리 시 정확도가 떨어질 수 있음

[표1] 매직 레이어스 주요 기능 요약

구분	작동 방식
레이어 분리 및 편집	• 이미지를 분석하여 구성 요소를 구분한 뒤, 각 요소를 개별적으로 편집할 수 있는 독립된 레이어로 분리
중첩 영역 자동 보정	• 구성 요소가 서로 겹쳐 가려진 영역을, AI가 주변 이미지 정보를 바탕으로 자동으로 채워 복원
텍스트 편집	• 이미지에 포함된 텍스트를 편집 가능한 문자 데이터로 변환하여, 폰트·크기·색상 등을 직접 편집할 수 있음

출처: 참고문헌 종합하여 재구성

[그림2] 캔바 편집 화면(좌), 매직 레이어 적용 후 레이어 분리 예시(우)



출처: Anna Wood, "Introducing Magic Layers, bringing creative control to AI content", Canva, 2026.03.11., <https://www.canva.com/newsroom/news/magic-layers>

매직 레이어스 도입에 따른 AI 이미지 편집 환경의 변화

- 편집 효율은 높아지나, AI 생성 여부를 판별하는 기준이 약화될 수 있다는 우려도 제기
 - AI는 이미지 생성 시 결과물을 단일 레이어로 출력하기 때문에, 레이어 단위의 선택적 편집이 어려워 일부만 수정하려면 프롬프트를 수정해 이미지 전체를 다시 생성해야 했음
 - 반면, 매직 레이어스를 활용하면 AI가 생성한 이미지도 구성 요소별로 분리하여 수정할 수 있어, 캔바는 이를 통해 반복적인 재생성에 드는 시간을 줄일 수 있을 것으로 전망함
 - 한편, AI가 생성한 이미지는 단일 레이어로 출력되기 때문에, 특정 작업물이 다층 레이어를 갖추고 있는지 여부가 창작자의 직접 제작을 확인하는 근거로 활용되기도 했음
 - 그러나 매직 레이어스가 레이어 분리 기능을 제공함에 따라 AI가 생성한 이미지도 다층 레이어 구조로 변환할 수 있게 되면서, 일각에서는 레이어 구조만으로 AI 생성 여부를 판단하기 어려워질 수 있다는 우려도 제기되고 있음¹⁾

참고문헌

- Anna Wood, "Introducing Magic Layers, bringing creative control to AI content", Canva, 2026.03.11., <https://www.canva.com/newsroom/news/magic-layers>
- Katelyn Chedraoui, "Canva May Have Just Cracked One of the Biggest Issues With AI Images", CNET, 2026.03.11., <https://www.cnet.com/tech/services-and-software/canva-ai-image-magic-layers-news/>
- Jess Weatherbed, "Canva's new editing tool adds layers to AI-generated designs", The Verge, 2026.03.11., <https://www.theverge.com/tech/893124/canva-ai-magic-layers-feature-beta>

1) Jess Weatherbed, "Canva's new editing tool adds layers to AI-generated designs", The Verge, <https://www.theverge.com/tech/893124/canva-ai-magic-layers-feature-beta>



저작권 이슈 브리프

SUMMARY

산업/기업

기술

중국 AI 만화극 시장 확대와 저작권 관리 체계 변화

AI 기반 중국 만화극 제작 산업의 성장

• 제작 공정 효율화를 통한 만화극 시장의 성장

- 중국의 AI 만화극(AI漫剧)은 만화·웹소설 원작을 영상으로 재창작한 콘텐츠로 생성형 AI(AIGC)*를 활용해 대본·캐릭터 등 제작 전 과정을 자동화하거나 반자동화하는 방식으로 제작됨
- 2025년 상반기 기준 중국 AI 만화극 시장의 월 평균 성장률은 92%를 기록하였으며, 연간 시장 규모가 200억 위안(원화 약 4조 1,900억 원)¹⁾을 초과할 것으로 전망되는 등 AI 도입 이후 시장이 급격히 성장함²⁾
- 특히 AI 도입으로 분당 제작 비용은 1,000~2,500위안(원화 약 21만 원~54만 원) 수준까지 절감되었으며, 기술 적용 방식에 따라 AIGC 2D·3D형부터 게임 편집기형까지 제작 유형이 다양해짐
- 그러나 기존 저작권법은 인간의 독창적 표현을 보호 대상으로 규정하고 있어, AI 산출물에 대한 이용자의 개입이 법적으로 '창작'의 영역에 해당하는지 판단하기 어렵다는 문제가 제기됨

* AIGC(AI-Generated Content): 생성형 인공지능 기술을 활용해 텍스트·이미지·음성·영상 등 콘텐츠를 자동으로 생성하는 방식으로, AI 만화극에서는 분경(컷) 설계, 캐릭터 생성, 음성 합성, 편집 등에 폭넓게 활용됨

• 콘텐츠 동질화·권리 분쟁 확산에 따른 기존 관리 체계의 한계

- AI 도입으로 제작 진입 장벽이 낮아지면서 유사한 소재와 스타일의 콘텐츠가 대량 유통되는 동질화 현상이 나타나고 있으며, 저품질 콘텐츠의 범람 문제도 함께 제기됨
- AI가 제작 전 과정에 깊이 개입하면서 창작자 중심의 기존 저작권 체계로는 권리 주체를 특정하기 어려워졌으며, 원저작물과의 저작권 분쟁 또한 증가하는 추세임
- 특히 AI 학습 데이터의 출처가 불분명한 경우가 많아, 기존 저작물과의 유사성을 사전에 확인하거나 차단하기 어려운 구조적 공백이 발생함
- 이에 대응하여 중국 정부는 AI 생성 콘텐츠에 대한 '명시적 식별 표시'를 의무화하고, 플랫폼 사업자가 침해 콘텐츠를 방치할 경우 공동 책임을 묻는 등 관리 감독의 강도를 높이고 있음
- 또한 중국 국가광전총국(国家广播电视总局)*은 2025년 11월 AI 만화극을 중점 심사 대상으로 명시하고, 기존의 사후 관리 방식에서 벗어난 선제적인 관리 체계 수립의 필요성을 공식화함

* 국가광전총국(国家广播电视总局): 중국의 방송·영상 콘텐츠를 관할하는 중앙 규제 기관, 드라마·영화·온라인 영상 등의 심사와 허가를 담당함

1) 1위안 = 209.51원(KEB 하나은행 매매기준율 적용, 2026.03.03.), 이하 동일 기준 적용

2) 李凤翔·田潇雨, "技术狂飙与发展红线:解析AI漫剧的合规密码", 星瀚微法苑, 2026.03.11., <https://www.36kr.com/p/32228342065494151>

AI 만화극에서 드러난 저작권·콘텐츠 관리의 쟁점

• AI 산출물의 저작물성 판단 기준과 제작 주체별 책임

- AI 산출물의 저작물성 인정 여부는 도구 활용을 넘어 인간 창작자의 ‘독창성 기여’가 개입되었는지가 핵심 판단 기준이며, 최근 중국 판례에서는 이를 보다 엄격하게 해석하는 추세임
- 이에 창작자는 프롬프트 전 과정과 수정 이력을 체계적으로 보관하고, 블록체인 기반의 디지털 저작권 식별자(DCI)* 등을 통한 선제적 권리 확보 노력이 요구됨
- 저작권 침해 시 플랫폼의 책임 소재는 생성 과정의 관여도와 실질적 이익 향유 여부에 따라 결정되며, 플랫폼이 템플릿 제공 등을 통해 제작에 깊이 관여할수록 직접 책임이 인정될 가능성이 높음

* 디지털 저작권 식별자(DCI, Digital Copyright Identifier): 중국에서 저작물의 권리 귀속 및 이용 이력을 식별·관리하기 위해 부여되는 고유 식별 체계로, 블록체인 등 기술을 활용하여 저작권 등록, 유통, 거래 과정의 추적과 증명 기능을 제공함

[표1] 중국의 AI 산출물 및 AI 플랫폼 저작권 관련 주요 판례

사건명	담당 법원	쟁점	결론
춘풍 사건(2023.11)	베이징 인터넷법원	AI 산출물의 저작권 인정 여부	저작권 인정
광저우 울트라맨 사건(2024.2)	광저우 인터넷법원	AI 플랫폼의 저작권 직접 침해 여부	직접 침해 인정
항저우 울트라맨 사건(2024.9)	항저우 인터넷법원	AI 플랫폼의 저작권 방조 침해 여부	방조 침해 인정
평씨 사건(2025.3)	수저우 장자강법원	AI 산출물의 저작권 인정 여부	저작권 불인정

출처: 李凤翔 田潇雨, “技术狂飙与发展红线: 解析AI漫剧的合规密码”, 星瀚律所, 2026.03.11., <https://36kr.com/p/3717903337009927>

• AI 산출물의 학습 데이터와 제작자의 의무

- 대다수의 AI 만화극 제작 도구가 학습 데이터 출처를 공개하지 않아, 제작자가 기존 저작물과의 유사성 여부를 사전에 확인하기 어려운 구조적 한계가 드러남
- 최근 중국 판례에서는 학습 데이터의 합법성 확인 의무를 최종 창작자와 AI 플랫폼 모두에게 폭넓게 인정하고 있으며, 이에 따라 AI 도구 선택 시 라이선스 조건에 대한 철저한 검증이 요구됨
- 특히 중요 프로젝트 진행 시에는 출처가 불분명한 외부 도구에 의존하기보다 오픈소스 모델을 직접 운용하여 안전한 데이터 위주로 제작 환경을 자체 관리하는 방식이 대안으로 제시됨
- 또한 외부 업체나 제3자로부터 공급받은 이미지·음악 등 AI 소재를 사용하더라도 저작권 문제가 발생하면 최종 제작자가 책임을 면하기 어려우므로 철저한 사전 검증이 요구됨

• 규제 당국의 분류 심사 및 식별 표시 의무화

- 중국 국가광전총국은 2024년 도입된 ‘분류 분층 심사’* 체계에 AI 만화극을 중점 관리 대상으로 편입하여, 투자 규모 및 소재에 따른 차등적 직접 심사를 강화함
- 2025년 9월 시행된 ‘AI 생성 합성 콘텐츠 표시 방법’에 따라, AI 산출물은 화면 내 명시적 표시와 파일 메타데이터 내 AIGC 식별 정보 삽입이 필수적이며 미이행 시 제재 대상이 됨
- 원작의 정신이나 캐릭터를 훼손하는 행위나 문화 정체성 침해 등은 엄격한 금지 기준으로 작동하며, 위반 시 플랫폼 차원의 삭제·차단 등 즉각적인 조치가 취해짐

* 분류 분층 심사: 투자 규모, 소재 등을 기준으로 콘텐츠를 중점·일반·기타 등 3개 등급으로 나누어 차등 심사하는 제도

시사점: 중국 사례로 본 AI 콘텐츠 제작·유통의 함의

• AI 창작 과정의 기록·데이터 관리의 필요성

- 콘텐츠의 저작물성을 판단하는 기준에 인간의 개입 여부도 포함됨에 따라, 창작 전 과정을 체계적으로 기록하고 보존하는 관리 방식이 필수적 요건으로 부상하고 있음
- 특히 AI가 학습에 사용한 데이터의 출처와 라이선스 계보를 확인하지 못할 경우 발생하는 소급 책임 리스크 방지를 위해, 기업 차원의 합법적 데이터셋 활용 및 검증 프로세스 구축이 필수적임
- 이는 개별 창작자의 노력을 넘어 AI 플랫폼이 창작 이력을 자동으로 보관하고 인증하는 저작권 관리 인프라를 서비스 내에 구축하는 노력이 병행되어야 함을 시사함

• 기술적 설계와 법적 책임이 결합된 체계로의 전환

- AI 만화극 산업에서 나타난 저작물성 판단·공급망 책임·콘텐츠 표시 의무의 결합은 장르의 문제를 넘어, AI가 제작 도구로 정착하는 모든 미디어 콘텐츠 산업이 직면할 구조적 과제를 선형적으로 보여줌
- 따라서 AI 플랫폼은 서비스 기획 단계부터 생성 과정의 관여 수준을 전략적으로 결정하여 리스크를 관리하는 '설계에 의한 책임(Liability by Design)' 전략을 검토할 필요가 있음
- 결국 중국의 사례는 사후 규제의 한계를 극복하기 위해 제작·유통 단계에 '심사·표시·계보 관리'를 내재화하는 모델이 산업의 법적 불확실성을 해소하는 하나의 실효적 방향이 될 수 있음을 보여줌

참고문헌

- 李凤翔·田潇雨, “技术狂飙与发展红线:解析AI漫剧的合规密码”, 星瀚律所, 2026.03.11., <https://36kr.com/p/3717903337009927>
- 电视剧司, “国家广播电视总局部署开展‘AI魔改’视频专项治理”, 国家广播电视总局, 2025.12.31., https://www.nrta.gov.cn/art/2025/12/31/art_114_72205.html
- 澎湃新闻, “AI漫剧快速成长, 各平台疯抢新赛道”, 2025.11.15., https://www.thepaper.cn/newsDetail_forward_31967160



주간 기술 동향

• 이미지 저작권 침해 입증의 한계와 의미론적 워터마킹을 통한 해결 방안

의미 요소별 파티션 설계를 통한 이미지 워터마킹 기술, SLICE

생성형 AI 활용 확산으로 저작권이 있는 이미지의 무단 사용이 급증하고 있음에도 여전히 창작자들이 자신의 저작물에 대한 저작권 침해를 입증하는 것은 쉽지 않다. 일례로, 2024년 미국에서 발생한 사진작가 조 지론(Joe Giron)과 핀테크 기업 라인 파이낸셜(Line Financial) 간의 소송은 이러한 현실을 단적으로 보여준다. 지론은 자신이 1993년 촬영한 미국의 록 가수 커트 코베인(Kurt Cobain)의 사진이 무단으로 사용된 사실을 확인하고 소송을 제기했으나, 법원은 그가 과거 라이선스 기록이나 시장 가격 증거를 제시하지 못했다는 이유로 손해배상 청구를 기각했다.

문제는 최근 디지털 이미지의 유통 구조가 복잡해지면서 침해 경로 추적 자체가 더욱 어려워졌다는 점이다. 지론의 사진은 이미지 제공 사이트를 거쳐 라인 파이낸셜에 도달했으나, 중간 유통 과정에서 원본 출처 정보는 완전히 소실되었다. 법원은 지론이 월페이퍼 사이트의 무단 게재를 알고도 조치하지 않았다는 점을 들어 침해 완화 노력의 부족을 지적했지만, 현실적으로 권리가 인터넷 전체를 감시하며 모든 무단 사용을 모니터링하는 것은 불가능하다. 이처럼 최근의 디지털 환경은 2차, 3차 유통을 거치며 원본 콘텐츠의 출처를 확인하기 어려워지고 있다.

이미지 워터마킹 기술은 이러한 문제를 해결할 수 있는 방안으로 주목받아 왔지만, 기존 방식들은 여전히 취약점을 안고 있다. 픽셀 기반 워터마크는 JPEG 압축이나 크롭 같은 간단한 편집에도 쉽게 손상되며, 초기 잠재 노이즈에 패턴을 삽입하는 방식은 생성형 AI를 활용한 워터마킹 위조에 취약하다. 최근 등장한 의미 기반 워터마크 기술은 이미지 내용과 워터마크를 결합하는 방식으로 취약점을 보완하였으나, 이미지 전체를 하나의 의미 코드로 묶는 구조를 가지고 있어 일부 요소만 자연스럽게 바꾸는 공격에는 무력하다는 사실이 밝혀졌다.

이에 따라 이미지 의미를 세분화하여 독립적 요소들로 분해하고, 각 요소를 공간적으로 구획화된 워터마크 파티션에 고정하는 접근이 제안되고 있다. 본 보고서에서는 구획화된 임베딩을 통한 의미론적 잠재 주입(Semantic Latent Injection via Compartment-alized Embedding, 이하 SLICE) 기술을 중심으로, 의미 요소별 파티션 설계가 어떻게 기존 방식의 한계점을 해결했는지 분석한다.

[사례] 구획화된 임베딩을 통한 의미론적 잠재 주입 워터마킹

• 기존 워터마킹 기술의 한계

- 전통적인 워터마킹 기술은 이미지의 픽셀이나 주파수 패턴에 보이지 않는 신호를 심는 방식이었으나, JPEG 압축이나 크롭(Crop, 자르기) 같은 간단한 편집만으로도 워터마크가 쉽게 손상됨
- 최근 AI 이미지 생성 기술이 발전하면서 트리링(Tree-Ring)*, 가우시안 셰이딩(Gaussian Shading)** 같은 새로운 워터마킹 기법이 등장했으나, 이들 역시 공격자가 생성형 AI 모델로 이미지를 재생성하면 워터마크가 제거되는 취약점이 있음
- SLICE는 이러한 한계를 극복하기 위해 이미지를 4가지 독립적인 의미 요소로 분해하고, 각각을 별도의 공간 영역에 고정하여 부분적 변조도 정확히 탐지할 수 있도록 설계됨

* 트리링(Tree-Ring): 확산 모델로 생성한 AI 이미지의 초기 노이즈 배열을 변환해 나이트 모양의 패턴을 삽입하는 워터마킹 기술

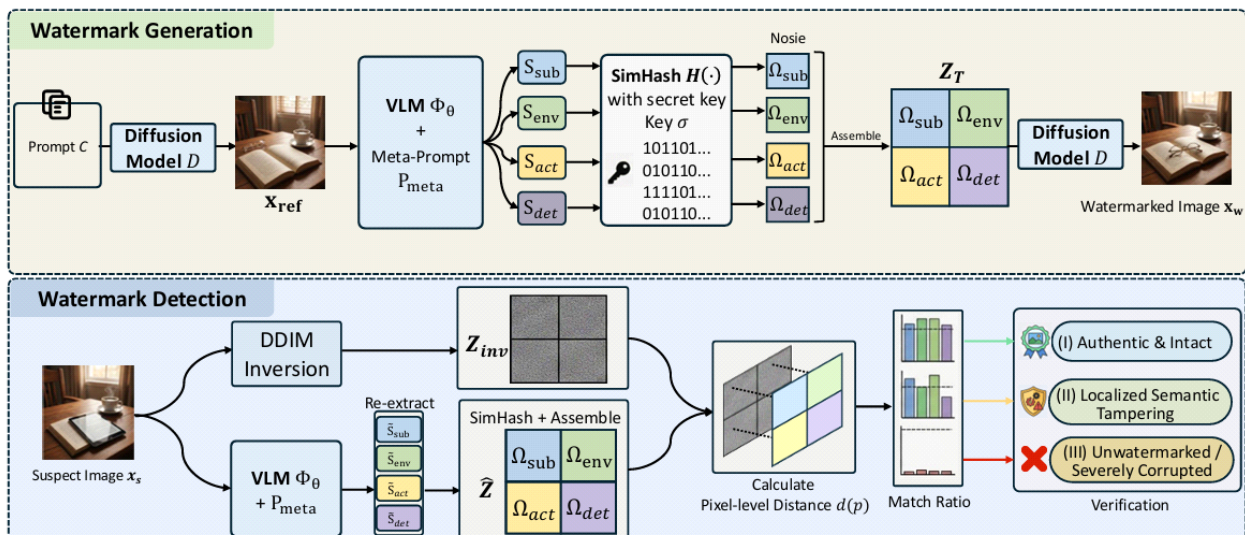
** 가우시안 셰이딩(Gaussian Shading): 확산모델 또는 3D 가우시안 스피클레팅 모델에서 사용되는 최신 비가시성 워터마킹 기술

• 이미지를 4조각으로 나누어 보호하는 원리

- SLICE는 먼저 비전 언어 모델(VLM)*을 사용하여 이미지를 분석함. 예를 들면, ‘전체 주제는 책, 배경은 창가 테이블, 상태는 정적인 배치, 세부사항은 안경과 커피잔, 햇빛’ 처럼 4가지 요소로 분해함
- 이렇게 추출한 각 의미 요소를 비밀 키로 암호화한 뒤, 이미지 생성의 출발점이 되는 초기 노이즈를 4개의 구역으로 나누어 각각 심음
- AI 이미지 생성 모델은 노이즈에서 이미지로 변환하는 과정을 수학적으로 역추적할 수 있기 때문에, 나중에 의심스러운 이미지를 다시 노이즈로 되돌린 뒤 각 구역의 비밀 키가 맞는지 확인할 수 있음
- 이 방식의 핵심은 4개 구역이 서로 독립적이라는 점임. 배경만 바뀌었다면 배경 구역의 비밀 키만 맞지 않고 나머지 3개는 그대로 유지되므로, ‘배경이 변조되었다’고 정확히 지목할 수 있음

* 비전 언어 모델(Vision Language Model, VLM): 컴퓨팅 비전 기술과 자연어 처리 기능을 결합하여 이미지와 텍스트를 동시에 이해하고 처리할 수 있는 인공지능 모델

[그림 1] SLICE의 작동 프로세스



출처: Zheng Gao와 6인, "SLICE: Semantic Latent Injection via Compartmentalized Embedding for Image Watermarking", arXiv, 2026.03.13., <https://arxiv.org/pdf/2603.12749>

• 3단계 진단 시스템으로 변조 여부와 위치 파악

- 검증할 이미지를 받으면 SLICE는 먼저 이를 노이즈 상태로 되돌리는 역변환 과정을 수행함. 그 다음 VLM으로 이미지에서 4가지 의미 요소를 다시 추출하고, 이를 바탕으로 노이즈를 재구성함
- 재구성한 노이즈와 실제 역변환한 노이즈를 4개 구역별로 비교하여 일치율을 계산함
- 모든 구역이 기준을 넘으면 State I(진본 및 무결)로 판정하여 원본으로 간주함. 이미지 전반에는 워터마크가 존재하지만, 일부 구역만 기준 미달이면 State II(국소 의미 변조)로 판정하여 실패한 구역을 통해 변조 위치를 특정함
- 만약 전반적으로 워터마크가 검출되지 않으면 State III(워터마크 없음 또는 심각한 손상)으로 판정함

• SLICE 성능 실험 결과

- 연구진은 가장 강력한 의미 편집 공격 기술로 평가되는 CSI*를 활용하여 SLICE를 포함한 주요 워터마크 기술 별 성능 실험을 진행함
- 실험 결과 CSI는 쉘(SEAL) 워터마크 기술에 대해 81%의 공격 성공률을 보였으며, 트리링 및 가우시안 셰이딩 기술을 상대로는 100%의 공격 성공률을 보임
- 반면, SLICE에 대해서는 CSI의 공격 성공률이 19%에 그쳐, SLICE가 기존 기술 월등한 방어 성능을 가진 것으로 나타남
- SLICE의 최대 장점 중 하나는 컴퓨팅 훈련이 필요 없는 방식이라는 것임. 기존 워터마킹 기술들은 AI 모델을 처음부터 다시 학습시켜야 해서 막대한 컴퓨팅 비용이 들었지만, SLICE는 이미 공개된 모델을 그대로 사용하기 때문에 상대적으로 소모되는 비용이 적음

* 일관성 보존 의미론적 주입(Coherence-Preserving Semantic Injection, CSI): 이미지 생성, 자연어 처리 등 AI 분야에서 원본 데이터의 핵심 의미나 흐름을 해치지 않으면서, 새로운 정보나 특징을 추가하거나 삽입하는 기술

결론 및 시사점

• 기술적 한계 및 개선 과제

- SLICE는 배경이나 인물 같은 일부 요소만 바꾸는 공격에는 강하지만, 이미지를 극단적으로 자르거나 크기를 변형하면 4개 구역의 배치가 무너져 워터마크 탐지율이 5%로 급락함. 이는 향후 크롭이나 회전 같은 기하학적 변형에도 견디는 기술 개발이 필요함을 보여줌
- 단기적으로는 SLICE 같은 기술과 전통적인 포렌식 기술을 함께 사용하는 것이 해법으로 제시됨
- 장기적으로는 크기 변형에도 강하면서 의미 편집도 탐지하는 차세대 기술이 필요하며, AI 생성 이미지와 기존 사진을 모두 보호하는 통합 시스템으로 발전하여 개인 창작자들이 실제로 사용할 수 있는 저작권 보호 수단이 될 전망이다

참고문헌

- Zheng Gao외 6인, "SLICE: Semantic Latent Injection via Compartmentalized Embedding for Image Watermarking", arXiv, 2026.03.13., <https://arxiv.org/pdf/2603.12749>
- Jonathan Bailey, "Kurt Cobain Photo Sparks Copyright Battle", Plagiarism Today, 2026.02.24., <https://www.plagiarismtoday.com/2026/02/24/kurt-cobain-photo-sparks-copyright-battle/>