

생성형 AI 영상 제작 기술 동향

한국저작권위원회
정보기술팀
문수연
2025. 9. 4.

보고서 요약

생성형 AI 영상 제작 기술은 인공지능이 텍스트·이미지·음성을 통해 자동 생성한 기술로, 영상 산업 전반에서의 새로운 가능성으로 주목 받고 있다.

생성형 AI 영상 제작 기술은 2014년 GAN의 등장으로 시작해 2017년 딥페이크로 대중에 알려졌고, 2021년 DALL·E 같은 모델로 영상 제작에 활용되기 시작했다. 이어 2022년 구글·메타의 텍스트-투-비디오 연구로 가능성이 확장되었으며, 2023년 Runway AI, OpenAI, 구글 등의 기업들이 플랫폼을 출시, 2024년 Sora와 Gemini 같은 고품질 모델이 등장하며 대중적으로 활용되기 시작했다.

이에 따라 다양한 산업 분야에서 비용 절감과 제작 효율성 향상, 창작 자유 확대라는 긍정적 효과를 제공하고 있다. 하지만 콘텐츠 제작 패러다임을 혁신하며 빠르게 확산되고 있는 만큼, 이용자 모두가 함께 책임 있는 활용과 윤리적 이용 자세가 필요하다.

1. 생성형 AI 영상 제작 기술의 개요

생성형 AI 영상 제작 기술은 인공지능이 텍스트·이미지·음성 등 다양한 입력을 기반으로 새로운 영상 콘텐츠를 자동으로 생성하는 기술을 의미한다. 사람이 직접 촬영하거나 편집하지 않아도, 원하는 장면과 스타일을 인공지능에게 지시하면 실제와 유사한 영상을 제작할 수 있다는 장점이 있다.

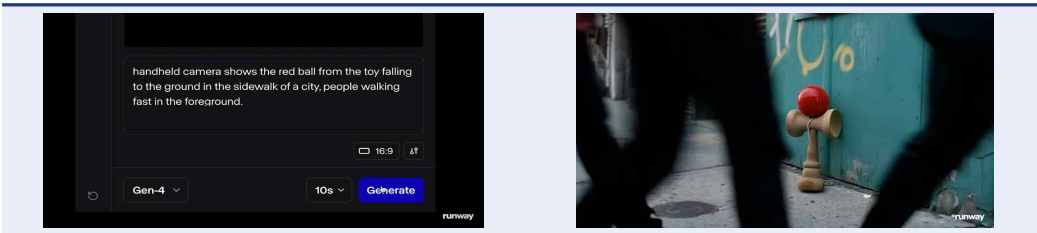
현재 생성형 AI 영상 제작 기술 분야에서는 스타트업부터 빅테크¹⁾까지 다양한 기업들이 폭넓게 참여하며 빠른 성장을 보여주고 있다. 챗GPT와 같은 텍스트 생성 AI나 미드저니(Midjourney)²⁾, 스테이블 디퓨전 (Stable Diffusion)³⁾ 등 이미지 생성 AI와 더불어 생성형 AI 분야의 한 축을 이루고 있다.

이렇듯 생성형 AI 영상 제작 기술의 급속한 발전에 따라, 영상 콘텐츠를 자동 생성하거나 편집할 수 있는 다양한 플랫폼들이 등장하고 있다. 대표적으로 Runway ML(런웨이 엠엘)의 Gen-4(젠-4), OpenAI의 Sora(소라), Google의 Veo 3(비오 3) 등은 사용자의 텍스트 입력만으로도 실제 촬영과 유사한 고품질 영상을 제작할 수 있는 기능을 제공하고 있다.

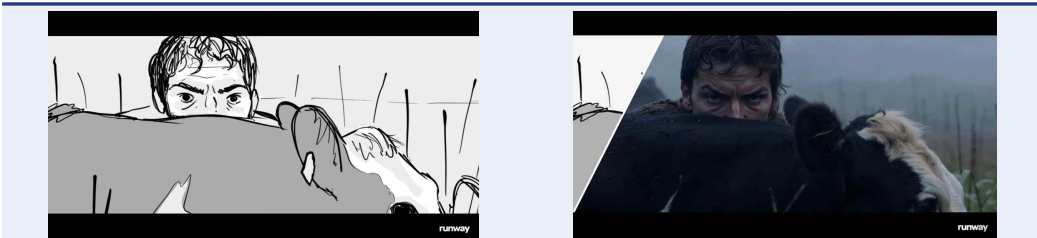
1) Runway AI의 Gen-4

인공지능 스타트업 Runway AI는 텍스트 또는 이미지를 입력하면 몇 초 길이의 영상을 만들어주는 텍스트-투-비디오 형식의 선두 주자로, 현재 제일 최신 모델인 Gen-4는 단순한 텍스트 지시 없이 시각적 참조를 활용해 스타일, 주제, 위치를 일관되게 유지하는 영상을 생성할 수 있다. 사용자는 원하는 장면의 구도를 설명하고, 인물이나 배경 이미지를 제공해 보다 정교한 영상을 제작할 수 있다.

| Gen-4에서 프롬프트를 통해 영상을 생성하는 모습



| Gen-4에서 스케치를 통해 영상을 생성하는 모습



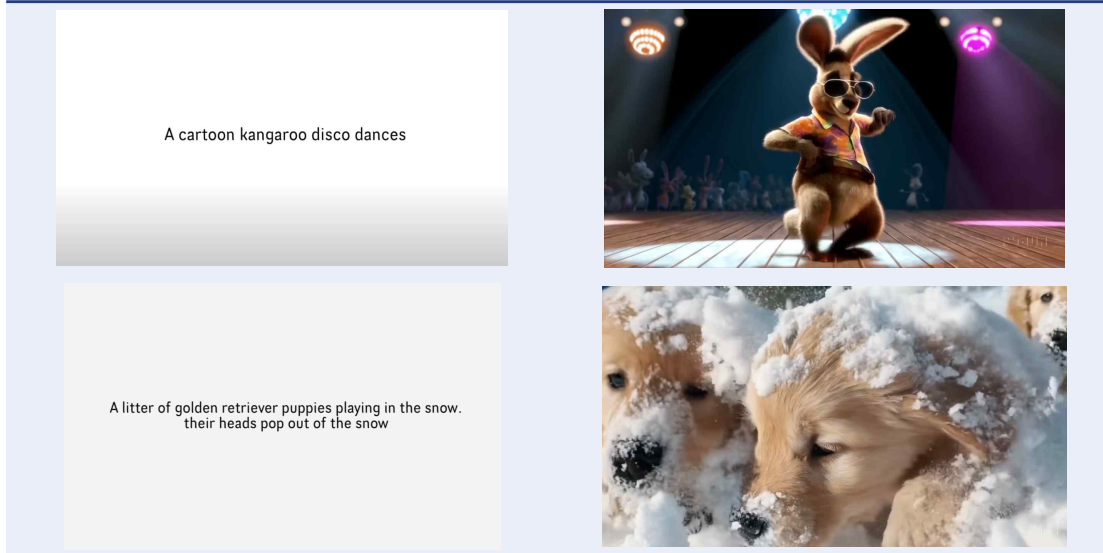
※출처: Runway 유튜브 채널, "Introducing Runway Gen-4 | Runway", 2025.3.31.

1) Big Tech 첨단 기술과 플랫폼 서비스 등을 기반으로 온라인상에서 다양한 서비스를 제공하는 대형 IT 기업
 2) 인공지능 연구소이자 해당 연구소에서 개발한 인공지능 소프트웨어
 3) 텍스트 설명에서 이미지를 생성하는 딥 러닝 모델

2) OpenAI의 Sora

챗GPT 개발사인 OpenAI가 공개한 최신 text-to-video(텍스트-투-비디오)⁴⁾ 모델의 Sora(소라)는 2024년 2월 시범 공개되며 큰 주목을 받았다. 입력 프롬프트- 맞춤화- 생성의 과정을 거쳐 고품질 영상을 생성하는 플랫폼으로서, 중력, 물의 흐름, 눈 쌓임 등 물리 법칙까지 자연스럽게 구현할 정도의 현실감을 자랑한다.

| Sora의 text-to-video(텍스트-투-비디오)를 통해 영상을 생성하는 모습



※출처: OpenAI 유튜브, "Introducing Sora - OpenAI's text-to-video model", 2024.2.17.

3) Google의 Veo 3

구글은 60초가량의 1080p 영상을 생성하는 Veo 1, 4K 해상도로 2분가량의 영상을 생성하는 Veo 2를 거쳐 현재의 Veo 3를 탄생시켰다. Veo 3는 음악 생성, 효과음 생성이 모두 가능할 뿐만 아니라 사운드 디자인 구현 수준이 매우 높아 서비스가 출시되기 이전부터 상당한 반응을 이끌어 내고 있다.

| Veo(비오) 3.0으로 생성된 영상



※출처: Google DeepMind 유튜브, "Veo 3 demo- Sailor and the sea", 2025.5.21.

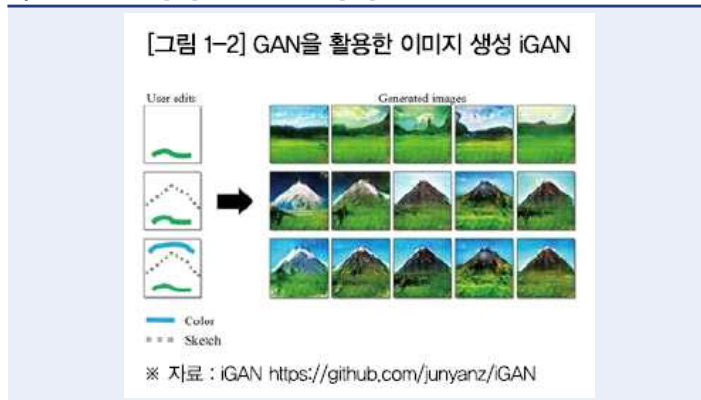
※출처: Google DeepMind 유튜브, "Veo 3 demo- Sweet typing", 2025.5.21.

4) 사용자가 입력한 텍스트 설명이나 지시에 따라 AI가 자동으로 동영상 콘텐츠를 생성하는 기술

2. 생성형 AI 영상 제작 기술의 발전 과정

생성형 AI 영상 제작 기술은 지난 10여 년간 기초 연구 단계에서부터 상용화, 대중적 활용으로 빠르게 진화해 왔다. 그 출발점은 2014년, 인공지능 연구자 이안 굿펠로우(Ian Goodfellow)가 제안한 생성적 적대 신경망(GAN, Generative Adversarial Network) 개념의 등장이다. 이는 인공지능이 새로운 이미지나 콘텐츠를 실제처럼 만들도록 하는 기반으로, 이후 생성형 AI 발전의 초석이 되었다.

| 적대 신경망 개념의 등장



※출처: 소프트웨어 정책연구소

2017년에는 GAN을 기반으로 한 딥페이크(Deepfake) 기술이 널리 알려지기 시작하며, 실제 사람의 얼굴을 정밀하게 합성하거나 입 모양을 바꾸는 기술이 대중에게 소개되었다. 이 시기부터 AI는 단순한 이미지 생성에서 나아가, 영상 콘텐츠의 구조와 표현까지 조작 가능한 단계로 넘어가기 시작했다.

2021년은 생성형 AI가 본격적으로 영상 콘텐츠 제작에 적용되기 시작한 해이다. OpenAI의 DALL·E를 비롯한 대규모 AI 모델의 등장으로 텍스트 입력만으로 이미지 생성이 가능해졌고, 일부 상업 영화나 광고 산업에서도 AI 기반 합성 기술이 실제 활용되기 시작했다. 이는 AI가 예술, 영상, 콘텐츠 산업에 본격 진입하는 신호탄이 쏘아 올려진 시점으로 평가된다.

이러한 흐름을 이어받아, 2022년에는 구글, 메타 등 빅테크 기업들이 텍스트-투-비디오(Text-to-Video)⁵⁾ 관련 연구 결과를 잇따라 발표하며 기술 가능성을 입증했다. 영상이라는 복잡한 매체를 AI가 직접 생성할 수 있다는 개념이 실험실 수준에서 점차 현실화되기 시작한 시점이었다.

2023년은 흔히 생성형 AI 영상의 상용화 원년으로 평가되며, 런웨이(Runway)와 같은 다양한 AI 영상 생성 플랫폼이 대중화되었고, 현대자동차의 AI 영상 광고와 같은 실제 브랜드 캠페인에 AI 영상이 활용되며 기술의 실용성이 입증된 해이기도 하다.

5) 텍스트 설명을 입력하면 AI가 이를 바탕으로 자동으로 동영상을 생성하는 기술

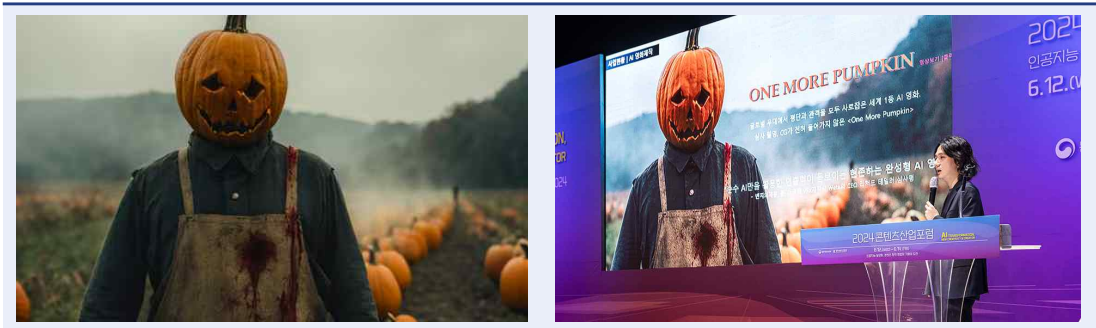
2024년은 생성형 AI 영상 제작 기술의 도약기로 평가되는 해이다. OpenAI가 발표한 '소라(Sora)', 구글의 '제미나이(Gemini)' 등 텍스트 기반의 고품질 영상 생성 기술이 공개되며, 사실적이고 생동감 있는 60초 이상의 고해상도 영상을 만들어내는 시대가 열렸고, 이와 함께 어도비(Adobe), 틱톡, 마이크로소프트 등 다양한 주류 플랫폼들도 AI 영상 생성 기능을 통합하며 생성형 AI 영상이 점차 디지털 콘텐츠 제작의 표준 도구로 자리 잡기 시작했다.

| 생성형 AI 영상 제작 기술의 발전 과정



위의 과정을 거친 생성형 AI 영상 제작 기술은, 콘텐츠 제작 효율성을 향상시키고 표현의 자유도를 높였으며, 궁극적으로 영상 제작의 접근성을 확대했다. 그중 영화 분야에서 <원 모어 펌킨>이 제작비 0원, 제작 기간 5일로 제작되어 주목받았으며, 이외에도 저예산으로도 특수효과를 구현하고 과거 작품을 복원하는 데 활용되고 있다.

| 생성형 AI 영상 제작 기술이 활용된 영화 <원 모어 펌킨>

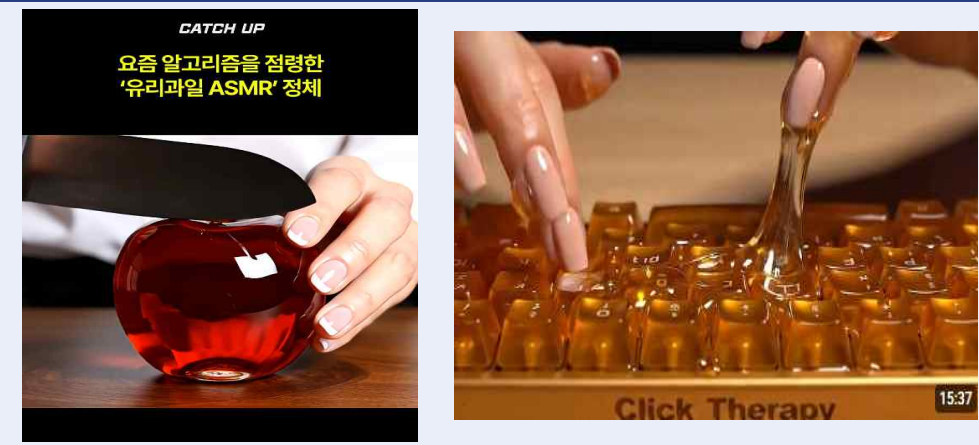


※ 출처: 한국영화데이터베이스, 한국콘텐츠진흥원

3. 생성형 AI 영상 제작 기술의 상용화

최근 유튜브 및 인스타그램 등 SNS에서도 생성형 AI 영상 제작 기술을 활용한 영상들의 비중이 크게 증가한 모습을 확인할 수 있다. 특히 유튜브에서 생성형 AI 영상이 빠르게 확산되며 새로운 콘텐츠 트렌드로 자리 잡게 되었고, AI 기반 콘텐츠 채널의 인기도 높아지면서 영상 시장에서 생성형 AI의 영상 제작 기술 영향력이 본격적으로 확대되고 있다. 텍스트나 이미지만 입력해도 짧고 흥미로운 영상을 제작할 수 있어 개인 크리에이터는 물론 기업의 광고 영역까지 활용 범위가 넓어지고 있다.

| 다양한 이용자들의 생성형 AI 영상 제작 기술 활용 사례



※출처: ccatch_up 유튜브

※출처: Click Therapy AI 유튜브



※출처: 정서불안김햄찌 유튜브

※출처: chef_Molty 유튜브

한편, 생성형 AI 영상 제작 기술은 저작권 침해와 딥페이크로 인한 사회적 혼란을 일으킬 수 있으므로, 출처와 이용 범위를 확인하고 관련 법률·가이드라인을 준수하며 책임감 있게 이용하는 태도를 지니는 것이 중요하다.

| 참고자료

- 소프트웨어 정책 연구소, “적대적 생성 신경망의 소개와 활용 현황”, 2017.09.29.
- 지역정보화, “AI 기술 발전에 따른 딥페이크의 두 얼굴”, 2024. Vol.141
- KCA, “딥페이크 기술과 투자 사기 광고”, 2024.01.02.
- 디지털투데이, “런웨이 동영상 생성 AI 모델 젠-공개... 사실적 움직임 구현”, 2025.04.01.
- KCA, “딥페이크 기술과 투자 사기 광고”, 2024.01.02.
- KOCCA N contents magazine, “AI 시대에도 창작의 주체는 사람입니다”
- ISC, “생산형 AI가 방송산업 및 인적자원에 미치는 영향”, 2023.06.27.