

특집논문 (Special Paper)

방송공학회논문지 제30권 제2호, 2025년 3월 (JBE Vol.30, No.2, March 2025)

<https://doi.org/10.5909/JBE.2025.30.2.142>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 생성형 AI 기반 단편영화 제작의 기술적 구현과 한계 분석

김채연<sup>a)</sup>, 주하은<sup>a)</sup>, 황진성<sup>a)</sup>, 유성권<sup>a)</sup>, 명윤호<sup>a)</sup>, 임영래<sup>a)</sup>, 인예원<sup>a)</sup>, 김다루<sup>a)</sup>,  
이승민<sup>b)</sup>, 박구만<sup>a)</sup>, 이영화<sup>c)</sup>, 양지희<sup>c)</sup>, 최영희<sup>b)</sup>†

## Technical Implementation and Limitation Analysis of Generative AI-Based Short Film Production

Chaeyeon Kim<sup>a)</sup>, Haeun Ju<sup>a)</sup>, Jinseong Hwang<sup>a)</sup>, Seongkwon Yoo<sup>a)</sup>, Yunho Myeong<sup>a)</sup>,  
Yeongrae Im<sup>a)</sup>, Yewon In<sup>a)</sup>, Daru Kim<sup>a)</sup>, Seungmin Lee<sup>b)</sup>, Gooman Park<sup>a)</sup>, Younghwa Lee<sup>c)</sup>,  
Jihee Yang<sup>c)</sup>, and Younghhee Choi<sup>b)</sup>†

### 요약

본 연구는 생성형 AI 기술을 활용하여 단편영화 제작의 효율성과 품질을 향상시키는 방안을 탐색하였다. 표정 연기, 동작 구현, 배경과 인물 합성의 세 가지 영역에서 연구를 진행하였으며, AILab Tools, Act-One, Flux Canny 모델, Remove.bg 등을 활용하여 감정 표현의 세밀함과 동작의 자연스러움을 개선하였다. 특히 표정 연기에는 Act-One을, 동작 구현에는 Flux Canny와 OpenPose를, 배경 합성에는 Remove.bg와 Stable Diffusion을 적용하여 기존 방식보다 품질을 높였다. 그 결과 제작 효율성이 증가하고 결과물의 품질이 개선되었으나, 감정 표현의 미세 조정과 데이터 의존성 등 기술적 한계가 확인되었다. 본 연구는 AI 기술의 가능성과 한계를 분석하고, 향후 다양한 제작 환경에서의 활용 가능성을 논의하였다.

### Abstract

This study explores the use of generative AI to enhance efficiency and quality in short film production. The research focuses on three key areas: facial expression refinement, motion implementation, and background-object integration. AILab Tools, Act-One, Flux Canny, and Remove.bg were employed to improve emotional subtlety and motion realism. Specifically, Act-One was used for facial expressions, Flux Canny and OpenPose for motion implementation, and Remove.bg and Stable Diffusion for background integration, resulting in higher production efficiency and improved visual quality. However, limitations such as challenges in fine-tuning emotions and reliance on training data were identified. This study analyzes the potential and limitations of AI in filmmaking and discusses its applicability to various production environments in the future.

Keyword : Generative AI, Short Film Production, Facial Expression Refinement, Motion Application, Background Integration

a) 서울과학기술대학교 스마트ICT융합공학과(Dept. of Smart ICT Convergence Engineering, Univ. of Seoul National University of Science and Technology)

b) 서울과학기술대학교 문예창작학과(Dept. of Creative Writing, Univ. of Seoul National University of Science and Technology)

c) 시그마케이주식회사(SigmaK Co., Ltd.)

† Corresponding Author : 최영희(Younghhee Choi)

E-mail: chanwch@seoultech.ac.kr

Tel: +82-2-970-6273

ORCID: <https://orcid.org/0009-0004-2286-5029>

※ 이 연구는 서울과학기술대학교 교내 일반과제 연구비 지원으로 수행되었습니다.

※ This research was supported by Seoul National University of Science and Technology.

※ 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 2024년 추계학술대회에서 발표한 바 있음.

· Manuscript January 14, 2025; Revised February 28, 2025; Accepted March 4, 2025.

Copyright © 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.

"This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered."

## I. 서 론

최근 몇 년간 생성형 AI 기술은 콘텐츠 제작 분야에서 새로운 가능성을 열며, 단순 보조 도구에서 창작 과정의 중심으로 발전하고 있다. 특히 영화 제작에서 AI는 기획, 촬영, 후반 작업 등 다양한 단계에서 활용될 수 있으며, 이는 제작 비용과 시간을 절감하면서도 창의적 결과물을 생성할 가능성을 제공한다. 실제로 영화 *sunspring*은 GPT-3 기반의 AI가 작성한 각본을 바탕으로 제작이 되었고 기존 작가들의 접근 방식과는 다른 방식의 서사 구조를 보여주었다. 또한 생성형 AI는 기존의 영상, 스토리, 음악 패턴을 분석한 후 이를 변형해 새로운 형태의 콘텐츠를 창출할 수 있다<sup>[1]</sup>. 그러나 현재 사용되는 AI 도구들은 감정 표현의 세밀함 부족, 움직임 구현의 부자연스러움, 배경과 인물의 조화 문제 등에서 기술적 한계를 드러내며, 영화 제작 과정에서 완성도를 저해할 수 있는 요인으로 작용한다.

예를 들어, 생성형 AI가 인물의 미세한 표정 변화를 생성할 때 감정의 맥락을 이해하지 못해 부자연스러운 움직임이 발생하며<sup>[2]</sup>, 배경과 인물 간 조화 역시 조명과 그림자 일관성 문제로 인해 어색하게 보일 수 있다<sup>[3]</sup>. 이러한 한계는 AI가 데이터 기반 패턴 분석을 통해 결과물을 생성하기 때문에 발생한다.

본 연구는 이러한 배경을 바탕으로, 생성형 AI를 활용한 단편영화 제작의 가능성과 한계를 분석하였다. 이를 위해 내부 제작 단편영화인 「The Face」<sup>[4]</sup>를 사례로 연구를 진행하였으며, 본 작품은 논문 연구를 위한 실험용 영화로 기획되었다.

「The Face」는 AI 기반 도구만을 활용하여 제작된 단편영화로, 기획·스토리 구성부터 배경과 인물 합성, 움직임 구현 까지 모든 과정을 AI와 협업하여 완성한 작품이다. 본 연구에서는 AI가 영화 제작 과정에서 창작 도구로서 어떤 역할을 수행할 수 있는지 평가하고, 기존 제작 방식과 비교하여 효율성과 기술적 한계를 검토하였다.

실제 제작 과정에서 AI가 일정 수준의 결과를 생성하는데 성공했으나, 세부적인 감정 표현의 한계, 복잡한 동작의 자연스러움 부족, 배경과 소품의 조화 문제 등 기술적 제약

이 확인되었다.

본 연구는 이러한 한계를 개선하기 위해 세 가지 주요 기법을 중심으로 접근하였다. 첫째, 표정 연기 기술을 통해 감정 표현의 세밀함을 높이고 자연스러운 캐릭터 묘사를 구현하고자 하였다. 둘째, 행동 입히기 및 애니메이션 적용 기법을 통해 복잡한 동작을 정확하고 자연스럽게 재현할 수 있는 방안을 모색하였다. 셋째, 배경과 인물의 합성 기법을 개선하여 AI 도구가 생성한 이미지와 영상의 일관성을 높이고, 보다 조화로운 장면을 구현하였다.

본 논문의 목적은 생성형 AI 기술이 단편영화 제작 과정에서 창의성과 생산성을 어떻게 증진할 수 있는지 탐구하는 것이다. 이를 위해 기존 AI 도구의 활용 사례와 한계를 분석하고, 제안된 개선 기법의 효과와 적용 가능성을 평가하였다. 연구 결과는 단편영화 제작에서 AI 기술의 실질적인 활용 가능성을 검토하며, 향후 기술 발전과 함께 다양한 장르와 조건으로 확장할 수 있는 방향성을 제시할 것이다.

## II. 기존 AI 영화 제작과 생성형 AI 영화 제작의 비교

기존 영화 제작은 창작자의 개입을 전제로 한 서사 구축, 연출, 연기, 편집 기술이 결합된 복합적인 예술 형태이다. 발터 벤야민(Walter Benjamin)은 영화가 기계적 복제 시대에 예술의 아우라(Aura)를 변화시킨다고 주장하였다<sup>[5]</sup>. 영화는 연극과 달리 동일한 경험을 반복적으로 제공할 수 있으며, 이를 통해 새로운 형태의 예술적 지각 방식을 창출한다. 기존 영화 제작에서는 감독과 배우, 촬영 감독 등의 협업이 중요하며, 실제하는 공간과 물리적 환경에서 촬영이 이루어져 사실성과 미학적 깊이를 더한다.

반면, 생성형 AI를 활용한 영화 제작은 기존 방식과 근본적으로 다른 패러다임을 제시한다. AI는 학습된 데이터를 기반으로, 자동으로 영상을 생성하며, 창작자의 개입 없이 특정 스타일이나 서사를 구현할 수 있다. 레프 마노비치(Lev Manovich)는 AI가 기존 미디어의 패러다임을 전환하며, 기존 요소들을 재조합하여 새로운 형식을 창출한다고

설명한다<sup>[6]</sup>. AI 기반 제작에서는 인물 애니메이션, 배경 합성 등 일부 과정이 자동화되며, 이는 제작 비용 절감과 시간 단축을 가능하게 하지만, 정교한 장면 연출이나 감정 표현에서 한계를 지닌다.

단편영화 제작 과정에서도 두 방식은 명확한 차이를 보인다. 기존 방식은 기획, 시나리오, 촬영, 편집 등의 단계를 거치지만, AI를 활용한 제작에서는 촬영과 후반 작업이 대체된다. 기존 방식에서는 창작자가 카메라 구도, 인물 연기, 배경 조성 등 모든 요소에 직접 개입하여 스토리를 구축하지만, AI 기반 제작에서는 자동 생성된 영상을 편집하는 방식으로 창작 과정이 변화한다. 기존 영화는 높은 제작비와 긴 제작 기간이 요구되지만, AI 기반 영화는 상대적으로 비용이 적고 장비 없이도 제작이 가능하다.

이러한 변화는 영화의 본질에도 영향을 미친다. 벤야민이 제시한 ‘아우라’ 개념에 따르면, 기계적 복제는 예술 작품의 독창성을 약화시킨다. AI가 만든 영상은 인간 창작자의 직접적인 개입 없이 제작되며, 이는 영화의 고유한 감성을 변화시키는 요인이 될 수 있다. 그러나 마노비치는 AI가 단순 복제가 아닌 새로운 조합과 변형을 통해 창조적인 결과물을 만들어낼 가능성을 강조한다. 질 베르 시몽동(Gilbert Simondon)의 기술 철학을 적용하면, 생성형 AI는 영화 제작 방식을 탈중앙화하고 창작 과정을 기술적 개입과 결합한 새로운 형태로 변환한다<sup>[7]</sup>. 창작자가 기계와 협력하는 방식이 전통적 영화 제작과 다르며, 이는 영화 제작의 주체성을 변화시키는 요소가 된다.

AI 기반 영화 제작은 또한 영화 산업의 접근성을 확대하는 역할을 한다. 기존 방식은 높은 기술적 진입 장벽과 제작 비가 요구되지만, AI를 활용하면 보다 많은 창작자가 영화 제작에 참여할 수 있다. 이는 영화 제작의 민주화를 촉진할 가능성이 있으나, 대량 생산된 콘텐츠로 인해 작품의 질이 저하될 가능성도 있다.

이처럼 생성형 AI는 영화 제작 방식을 혁신하지만, 예술적 독창성과 감성 표현의 한계라는 도전 과제도 함께 내포하고 있다. 향후 연구에서는 이러한 기술이 기존 영화 제작 방식과 조화를 이루며 창작의 새로운 가능성을 제공할 수

있는 방안을 모색해야 할 것이다.

### III. 생성형 AI 기반 단편영화 제작 워크플로우 및 한계 분석

#### 1. 상용 도구를 활용한 단편영화 제작

영화 제작에서는 스토리와 메시지 전달이 중요한 요소로, 영화 시나리오는 인간 창작자가 작성하고 이를 기반으로 생성형 인공지능 도구를 활용해 영상을 제작하였다. 인물의 목소리는 Eleven Labs<sup>[8]</sup>의 Text-to-Speech 기능으로 생성되었으며, 효과음은 저작권 문제가 없는 음원과 AI 도구를 병행하여 사용하였다. 생성된 짧은 영상들은 Adobe Premiere Pro<sup>[9]</sup>로 편집하여 완성도를 높였다.

영상 제작 시 일관된 시각적 표현을 유지하기 위해 시드 번호 고정 기술을 활용하였다. 이를 통해 등장인물의 특성, 화면의 색감, 스타일 등을 통일감 있게 유지할 수 있었으며, 협업 과정에서도 동일한 시각적 기준을 공유할 수 있었다.

이미지를 기반으로 움직임을 추가하는 Image to Video 방식은 세부적인 시각 표현을 구현하는 데 유리했으며, Midjourney<sup>[10]</sup>의 다양한 기능을 통해 입력 이미지를 생성하고 기존 콘텐츠를 변형하거나 새로운 콘텐츠를 제작할 수 있었다. 반면, Text to Video 방식은 텍스트를 기반으로 영상을 생성하는 기술로, 세부 제어는 어렵지만 창의적인 결과물을 얻는 데 적합하였다.

또한, 카메라 구도와 초점거리, 카메라 움직임 등을 명령어로 지정하여 의도한 연출을 구현할 수 있었으며, 이를 통해 영상의 분위기와 스토리 전개를 효과적으로 전달하고 관객의 몰입도를 높일 수 있었다.

이러한 AI 도구를 활용한 제작 과정은 기존 영화 제작 방식과는 다소 차이가 있으며, 텍스트 기반 시나리오 구상에서부터 이미지 생성, 영상 합성, 편집 및 후반 작업까지 AI의 개입이 이루어진다. 본 연구에서는 이러한 과정을 체계적으로 정리하였으며, 그림 1은 AI 도구를 활용한 단편영화 제작의 주요 단계를 시각적으로 나타낸 것이다.

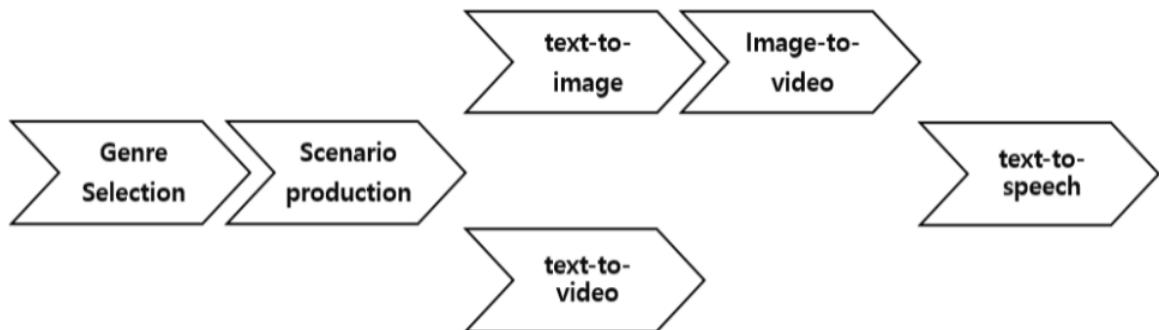


그림 1. 상용 도구를 활용한 단편영화 제작의 순서도  
Fig. 1. Flowchart of short film production using commercial tools

## 2. 상용 도구를 활용한 단편영화 제작의 한계점

### 2.1 표정 연기 기술 관점에서의 한계

영화 제작에서 인물의 감정 표현은 스토리 전달과 관객 몰입에 중요한 역할을 한다. 그러나 본 연구에서 사용한 생성형 AI 도구들은 기쁨, 슬픔, 화남과 같은 대표적인 감정 이외의 세밀한 감정을 나타내지 못하였다. 프롬프트 기반으로 미묘한 감정을 정확하게 구현하는 데 한계가 있었다.

예를 들어, “창피하다”나 “쑥스럽다”와 같은 감정을 직접 프롬프트로 입력했을 때, AI는 과장된 표정이나 무표정을 생성한다. 이러한 문제로 인해 제작자는 “인물이 땅을 본다”와 같이 행동을 통해 간접적으로 감정을 표현해야 했으며, 이는 원하는 감정의 섬세한 전달에 어려움을 초래했다.

또한, 여러 인물이 등장하는 장면에서는 각 인물의 고유한 감정을 구분하여 표현하기 어려웠다. 결과적으로, 감정 선이 중요한 장면일수록 표정 표현의 일관성과 자연스러움이 부족하여 영화의 완성도가 저하되는 문제가 발생하였다.

### 2.2 행동 및 애니메이션 적용 관점에서의 한계

인물의 동작이나 여러 대상 간 상호작용을 표현할 때, AI가 무작위로 형상을 왜곡하거나 비현실적인 자세를 생성하는 문제가 관찰되었다. 예를 들어, 주먹으로 상대방의 얼굴을 치는 장면을 프롬프팅 했으나, 주먹이 움직이지 않고 좌우로만 흔들리는 등, 의도하지 않은 움직임이 생성되기도 하였다.

또한, 인공지능을 통해 영상을 생성하면, 인물이 움직이면서 인물의 모습이 변하기도 하고, 카메라 구도가 바뀌며

배경이 변화하기도 하여 영상 안에서 일관성 유지가 어렵다. 특히 영상 내에서 움직임의 범위가 클 때, 영상 속 피사체들의 형태 유지가 어렵다는 문제점이 있었다.

더욱이, 영상의 디테일 부분은 생성할 수 없는 경우가 많았다. 예를 들어, 손으로 넥타이를 매는 장면을 생성할 때, 손 모양과 움직임을 텍스트로 입력하면 이에 대한 생성을 하지 못한다. 특히 두 명 이상의 인물이 동시에 움직이거나 사물과 상호작용을 하는 장면을 구현하려 하면, 텍스트 프롬프트만으로는 동작을 일일이 지정하기가 쉽지 않았다. 이 때문에, 원하는 액션을 구체적으로 입력하더라도 AI가 이를 정확히 이해하지 못해 생성된 결과물이 의도와 다르게 왜곡되거나 상호작용이 부자연스러운 경우가 발생했다.

### 2.3 배경과 인물의 합성 관점에서의 한계

Midjourney의 이미지 부분 편집 기능을 사용할 때, 참고 이미지를 넣고 부분 편집 명령어를 적용하고자 하였으나, 실제로는 기능이 발휘되지 않는 문제가 있었다. 예를 들어, 그림 2에서 보는 바와 같이 파란 넥타이를 맨 남성의 모습에서 넥타이의 색상을 다른 색상으로 바꾸기 위해 부분 편집 기능을 사용했으나, 이전 이미지를 참고하여 부분 편집이 적용되어 그대로 파란 넥타이가 나왔다. 즉, 다른 색상으로 바꿔주지 않고 파란 넥타이를 맨 남성의 모습을 다시 반복해서 생성했다.

특히, 그림 같은 연출이 지속적으로 발생하면, 예술적 측면에서는 독특한 효과를 낼 수 있을지 모르지만, 리얼리즘을 중시하는 장면에서는 몰입도를 해치게 된다. 실사 영상을 요구하는 프롬프트를 사용했어도 ‘fantasy’, ‘mystery’,

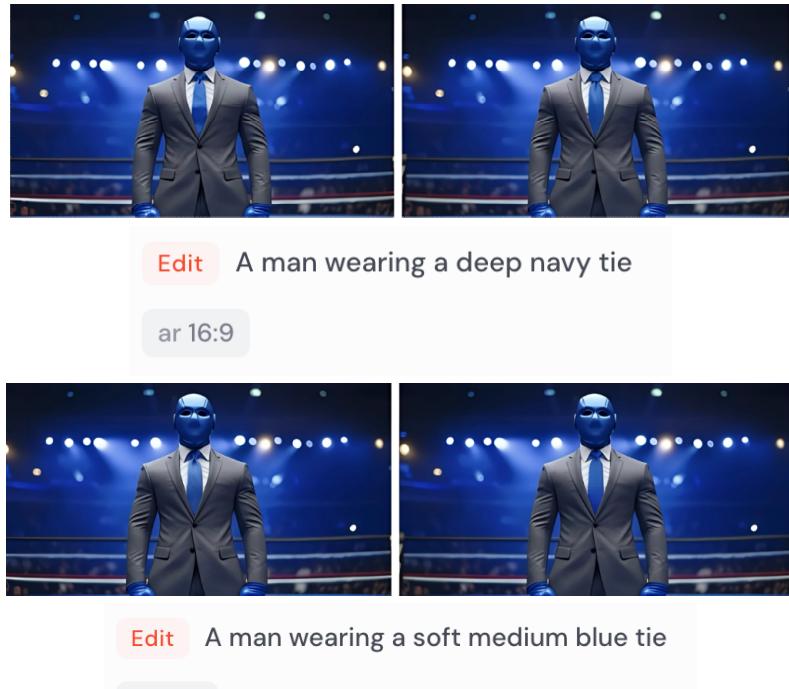


그림 2. 이전 이미지를 반복하여 생성하는 Midjourney  
Fig. 2. Midjourney Repeating the Same Image Generation

‘dreamlike’ 같은 공상적인 분위기의 영상을 생성하는 명령어를 입력하면 영상의 분위기나 색감을 조정하여 분위기를 조정하기보다 영상 스타일을 바꿔 그림 같은 영상을 생성한다.

#### IV. 생성형 AI를 활용한 영상 제작 기법의 고도화

##### 1. 화면 구성 요소별 생성 및 조합 기법

본 연구에서는 「The Face」 제작 과정에서 기존에 활용해온 Runway, Midjourney 등 다양한 생성형 AI 도구의 한계를 발견하였다. 이러한 문제점을 해결하기 위해 세 가지 핵심 기법을 선정하고 심층 조사를 수행하였으며, 본 장에서는 해당 기법들의 이론적 배경과 적용 과정을 구체적으로 제시하고자 한다.

###### 1.1 표정 연기 기술

첫 번째로 제안하는 개선된 영화 제작 기법은 표정 연기

기술이다. 일반적으로 원하는 표정을 정확히 표현하기 위해 생성형 AI를 활용하지만, 프롬프트 작성 후 결과물을 생성하는 과정에서 원하는 표정이 너무 과하게 구현되거나 제대로 구현되지 않는 문제가 빈번히 발생한다. 이러한 문제를 해결하기 위해, Pretrained Model과 상용 도구를 함께 활용하는 방법을 제안한다.

대표적인 Pretrained Model로는 GAN(Generative Adversarial Network)<sup>[11]</sup>이 있다. 특히 cGAN(Conditional GAN) 기술을 이용하면 입력 이미지와 조건(예: 특정 표정)을 동시에 학습하여, 원하는 표정을 상대적으로 자연스럽게 조작할 수 있다<sup>[12]</sup>. 또한 StyleGAN을 적용하면 표정 외에도 얼굴 윤곽, 머리카락 스타일 등 세밀한 부분까지 조정이 가능해 더욱 다양한 시각적 효과를 구현할 수 있다<sup>[13]</sup>.

상용 도구로는 AILab Tools<sup>[14]</sup>가 대표적이다. 업로드한 이미지의 얼굴을 분석하여 표정을 바꿀 수 있는 상용 도구이며 12가지의 표정 프리셋에 맞게 이미지를 변경할 수 있다.



그림 3. Act-One을 활용한 표정 변경 화면 예시 (기존 이미지/Ace-One을 활용한 생성 이미지)

Fig. 3. Example of Facial Expression Modification Screen Using Act-One (Original Image/Generated Image using Act-One)

이 도구를 활용하면 한 번의 클릭만으로 여러 표정을 손쉽게 바꿀 수 있는 장점이 있으며 복수의 감정 표현을 추가하여 다양한 연출을 시도할 수 있다.

다른 상용 도구로 Runway Gen-3의 Act-One 기능이 있다. Act-One 기능을 사용하면 자신이 업로드한 영상을 기반으로 영상을 생성할 수 있다. 원하는 표정이 제대로 구현되지 않는다면 자신이 직접 원하는 표정을 찍어서 생성할 수도 있다.

Act-One을 사용하면 프리셋에 없는 다양한 표정을 자신이 원하는 대로 설정할 수 있는 장점이 있다.

이때 Pretrained Model은 정교한 학습 데이터셋을 바탕으로 한 고품질 표정 생성을 제공하고, 상용 도구는 사용자 친화적 인터페이스를 통해 직관적이고 빠른 표정 조작을 지원한다. 결과적으로 이러한 방법을 통해 캐릭터의 감정과 스토리를 더욱 효과적으로 전달할 수 있을 것으로 기대된다.

## 1.2 행동 입히기 및 애니메이션 적용 기법

영화 제작에 인공지능을 활용할 때, 텍스트 프롬프트만으로 인물의 동작을 정확하게 표현하기는 쉽지 않다. 손·발처럼 복잡한 부위가 왜곡되거나, 전혀 의도하지 않은 포즈가 생성되는 문제가 빈번히 발생하기 때문이다. 이를 해결하고자 본 연구에서는 Hunyuan Video Model<sup>[15]</sup>과 Flux Tools<sup>[16]</sup> + OpenPose<sup>[17]</sup>를 사용한 행동 입히기 기법을 제안한다.

우선 오픈 소스로 공개된 Hunyuan Video Model은 인물

이미지와 Skeleton 데이터를 합성해, 춤과 같은 세부 동작을 정교하게 구현할 수 있다. 기존 Runway로 영상을 생성할 때는 텍스트만으로 인물의 움직임을 지정해야 하므로 의도한 동작을 정확히 구현하기 어려웠지만, 특정 동작을 담은 영상 이미지에서 추출한 Skeleton을 그대로 복사해 적용함으로써 정확한 움직임을 효과적으로 재현할 수 있게 되었다. 그림 4는 본 모델을 활용한 예시로, 원본 인물 이미지, 추출된 Skeleton, 합성 후 동작이 변형된 결과물을 보여준다.

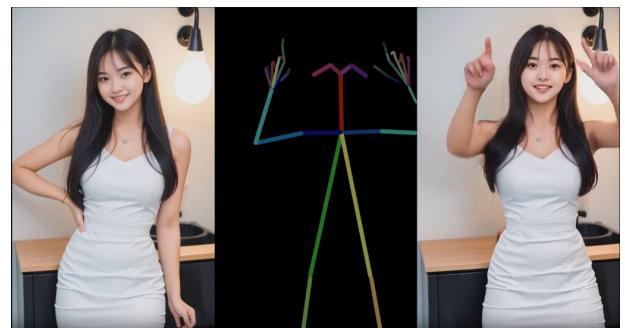
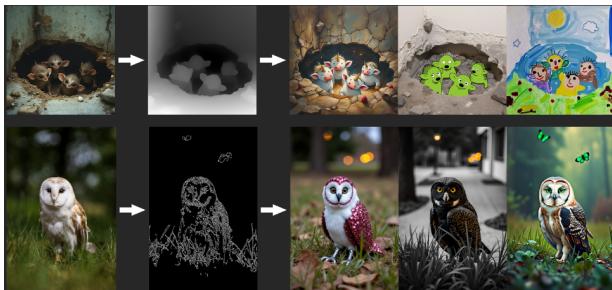


그림 4. Hunyuan Video Model을 활용한 Skeleton 합성 예시<sup>[18]</sup>

Fig. 4. Example of Skeleton Synthesis Using Hunyuan Video Model<sup>[18]</sup>

또 다른 접근으로, Flux Tools와 OpenPose를 결합하는 방법이 있다. Black Forest에서 발표한 Flux Tools는 이미지 생성 시 외곽선(Canny), 깊이 정보(Depth) 등을 활용하여 창작자가 원하는 동작을 정확히 재생성할 수 있는 모델이다. 기존 Runway나 Midjourney 등에서 손·발처럼 복

잡한 형상이 왜곡되거나 의도치 않은 포즈가 생성되는 문제를, Flux Tools가 정확한 라인 정보를 인식해 보완해 준다.

그림 5. Flux Tools를 활용한 이미지 변환 예시<sup>[19]</sup>Fig. 5. Example of Image Transformation Using Flux Tools<sup>[19]</sup>

또한 OpenPose와 결합하면, OpenPose가 만들어낸 Skeleton(자세 정보를 담은 관절 데이터)을 Flux Tools에 입력하여 정확한 동작이 반영된 이미지를 합성할 수 있다. 즉, 창작자가 직접 원하는 자세를 지정한 뒤 해당 포즈에 맞춰 인물·배경·소품 등을 새로 생성·보정함으로써 높은 품질의 합성 이미지를 얻을 수 있다는 장점이 있다.

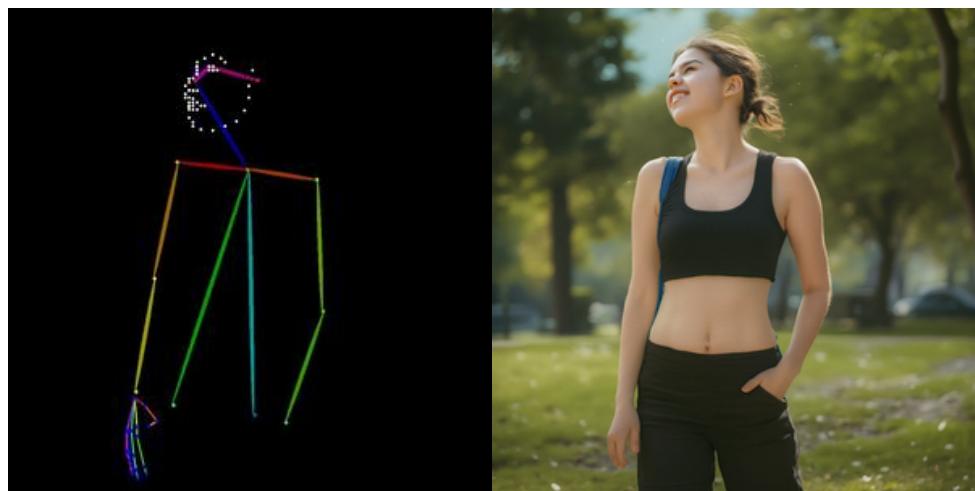
### 1.3 배경과 인물의 합성 기법

다음으로 제안하는 영화 제작 기법은 배경과 인물의 합

성이다. 일반적으로 생성형 AI를 통해 영상을 제작할 때, 창작자가 구상한 대로 결과물이 구현되면 이상적이겠지만, 도구화된 AI를 활용하는 경우 예상과 다른 결과물이 빈번히 발생한다. 이를 해결하기 위해, 본 연구에서는 Pretrained Model과 상용 도구를 동시에 활용하는 방안을 제안한다.

대표적인 Pretrained Model로는 Stable Diffusion<sup>[21]</sup>이 있으며, 그중 Inpainting 기술을 사용하면 이미지 내 특정 영역을 수정하거나, 새로운 요소를 추가할 수 있다. 예를 들어, 인물의 의상·머리 스타일·배경 등 창작자가 변경하고자 하는 부분을 손쉽게 교체·보완함으로써 의도한 장면을 구현할 수 있을 것으로 기대된다. 특히 실사 기반 영상에서 소품이나 배경이 부족하거나, 인물 표현에 오류가 생긴 경우에도 추가 촬영 없이 Inpainting만으로 완성도 높은 결과물을 얻을 수 있다는 점이 장점이다.

Pretrained Model이 아닌 상용 도구인 Remove.bg<sup>[22]</sup>를 사용할 때도, 전반적인 적용 방식은 유사하다. Stable Diffusion의 Inpainting과 유사한 브러쉬 기능을 활용해 원하는 부분이나 배경 등을 제거·수정한 뒤, 창작자가 원하는 이미지를 합성하는 방식이다. 이 과정에서 창작자는 직접 이미지를 선별하여 합성할 수 있으므로, 예전대 유명 예술 작품이나 특정 사진을 삽입해 완성도 높은 장면을 연출할 수 있다.

그림 6. OpenPose Skeleton 적용 예시<sup>[20]</sup>Fig. 6. Example of OpenPose Skeleton Application<sup>[20]</sup>

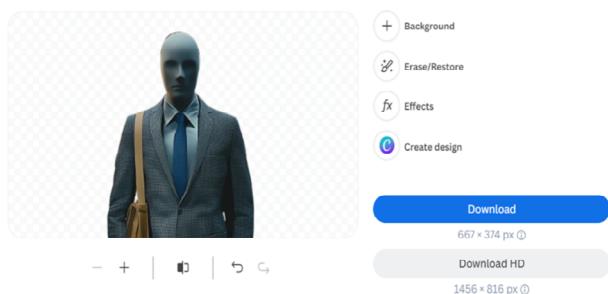


그림 7. Remove.bg를 활용한 배경 제거 및 인물 합성 예시  
Fig 7. Example of Background Removal and Person Composition Using Remove.bg

다만, 실제 사용 결과 AI 스스로도 원하는 결과물과 유사한 장면을 구현하고 인물과 배경을 자연스럽게 합성해주는 경우가 많았다는 점에서, 반드시 이와 같은 보조 도구를 활용하지 않아도 충분히 만족스러운 결과가 나오는 상황이 있다. 그럼에도 불구하고, 특정 상황이나 보다 세밀한 조정이 필요한 경우에는 앞서 제안한 기법들을 결합해 더욱 효과적인 결과물을 제작할 수 있을 것으로 판단된다.



그림 8. AI Lab tool을 활용한 표정 생성 예시 (기존 이미지/생성 이미지)  
Fig. 8. Example of Facial Expression Generation Using AI Lab Tool (Original Image/Generated Image)

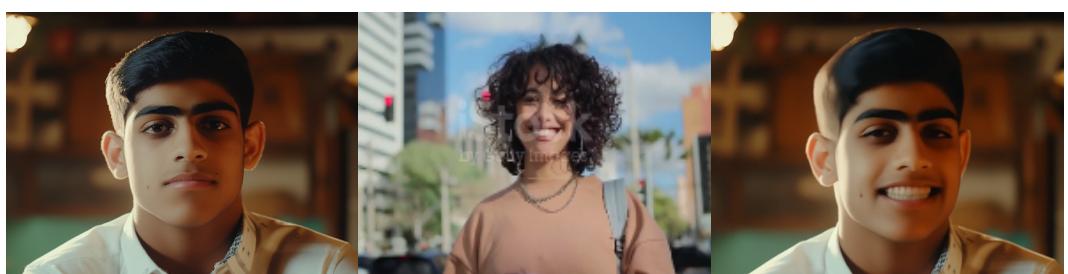


그림 9. Act-One를 활용한 표정 변경 화면 예시 (기존 이미지/레퍼런스 영상/생성 영상)  
Fig. 9. Example of Facial Expression Modification Using Act-One (Original Image/Reference Video/Generated Video)

## 2. 화면 구성 요소별 생성에 의한 화면 개선

### 2.1 표정 연기 기술

본 연구에서는 기존에 부자연스럽게 생성된 표정을 자연스럽게 구현하기 위해 제안된 표정 연기 기법을 활용하여 영상을 재구성하였다. 특히, 무표정이었던 인물의 얼굴을 뿌듯한 감정을 강조하는 표정으로 변경하는 작업을 수행하였다.

AILab Tools를 사용한 경우, 클릭 한 번으로 간편하게 웃는 표정을 생성할 수 있어 사용이 매우 편리하였다. 그러나 프리셋에 맞추어 변형만 가능했기 때문에, 미묘한 감정 표현에는 한계가 있었다. 또한, 툴이 이전의 맥락을 고려하지 못하므로, AI 영화를 제작하는 과정에서 원하는 표정을 정확히 연기 어려운 문제가 존재하였다.

반면, Runway Gen 3의 Act-One 기능을 활용한 경우, 레퍼런스 영상을 통해 매우 유사한 표정을 생성할 수 있었으며, 자연스러운 표정의 영상을 즉각적으로 얻을 수 있는 장점이 있었다. 그러나 원하는 장면에 적합한 레퍼런스 영상을 찾는 데 상당한 시간이 소요되며 레퍼런스 영상의 선택

이 제한적일 경우 AI 영화 제작에 효과적으로 활용하기 어려운 점이 있었다.

## 2.2 행동 입히기 및 애니메이션 적용 기법

다음으로, 인공지능을 활용해 인물의 동작을 자연스럽게 표현하고, 실시간 애니메이션을 구현하는 방안에 대해 살펴보자 한다. 본 연구에서는 이러한 기법의 예시로

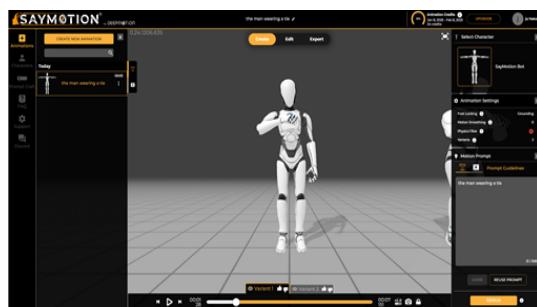


그림 10. DeepMotion을 활용한 캐릭터 동작 생성 예시  
Fig. 10. Example of Character Motion Generation Using DeepMotion

DeepMotion<sup>[23]</sup>을 적용하여, 손·팔 등 신체 움직임부터 표정·행동까지 효과적으로 입히는 가능성을 확인하였다.

먼저, Runway로 넥타이를 매는 남성을 생성했을 때 손 모양이나 손가락이 기괴하게 변형되는 문제가 빈번히 발생하였으나, DeepMotion을 사용한 결과 손·팔 움직임이 비교적 자연스럽게 표현되었다. 특히 영상 입력을 기반으로 Skeleton을 생성·적용하는 방식은 손가락 각도나 개수 등 세부 디테일을 보다 정확히 구현할 수 있을 것으로 기대된다.

또한 Runway에서 구현하기 어려웠던 복싱 장면을 DeepMotion으로 제작해 본 결과, 팔·다리 움직임이 훨씬 자연스러워졌고, 손가락까지 세밀하게 묘사된다는 점이 확인되었다. Runway로 복싱 장면을 구현하였을 때는 영상이 진행이 될 때 사람의 손가락의 모양, 관절의 위치가 부자연스럽게 왜곡이 되는 단점이 있었지만, Skeleton 기반 기법을 적용한 DeepMotion으로 생성을 한 경우 영상이 진행되어도 사람의 관절이나, 손가락 개수 등이 왜곡되지 않는다는 장점이 있었다.



그림 11. 복싱 장면 구현 예시 (Runway/DeepMotion)  
Fig. 11. Example of Boxing Scene Implementation (Runway/DeepMotion)

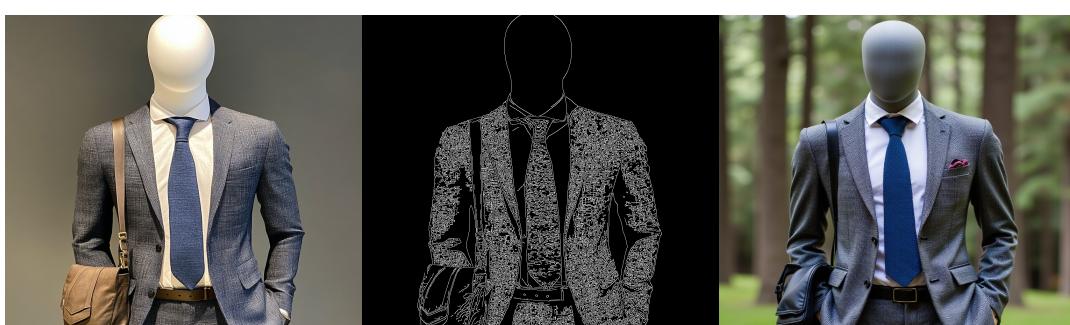


그림 12. Flux Canny 모델을 활용한 단계별 합성 예시  
Fig. 12. Step-by-Step Composition Example Using the Flux Canny Model

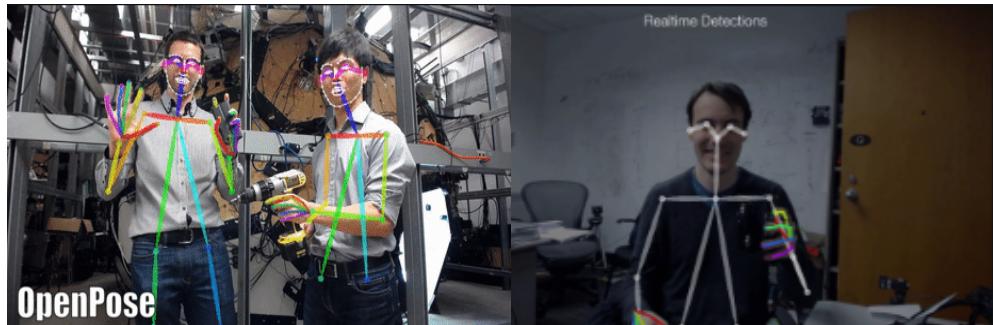


그림 13. Openpose를 활용한 실시간 자세 인식과 Skeleton 정보를 적용한 모습<sup>[25]</sup>

Fig. 13. Real-Time Pose Recognition and Application of Skeleton Information Using OpenPose<sup>[25]</sup>

기존에 Runway나 Midjourney에서 인물의 동작을 구현할 경우, 손·발과 같은 세부 디테일이 낮은 품질로 생성되어 어색함이 발생하는 문제가 있었다. 이를 개선하기 위해, Flux Canny 모델을 활용하여 외곽선(Canny Edge) 정보를 정확히 인식하고 재생성함으로써 이미지 품질을 향상하는 방안을 시도하였다.

추가적으로, ComfyUI<sup>[24]</sup> 툴을 통해 OpenPose와 결합하면 창작자가 원하는 구체적 동작을 구현하는 것도 가능해진다. 즉, OpenPose로 인물의 자세(Skeleton)를 지정하고, Flux Canny 모델로 외곽선과 배경 정보를 재생성하여 정교하면서도 원하는 동작을 구현할 수 있다는 결론에 도달하였다.

한편, Adobe Character Animator<sup>[26]</sup>는 Pose Estimation을 통해 추출한 Skeleton 정보를 활용함으로써 캐릭터에 움직임과 표정을 부여할 수 있는 툴이다. 정적 이미지를 실시간

으로 애니메이션화할 수 있다는 장점이 있지만, 2D 애니메이션에 최적화되어 있어 실사화된 인공지능 영화 제작에는 한계가 있다. 그럼에도 간단한 움직임과 표정만으로 짧은 2D 콘텐츠를 빠르게 제작해야 할 때에는 충분히 유용한 솔루션으로 평가된다.

위의 사례들을 통해 행동 입히기 기술을 적용하여 영상을 제작한 사례가 동작의 자연스러움, 자세의 정확성, 애니메이션 품질과 같은 영상 평가 기준에서 AI 도구만 이용한 영상 생성보다 더욱 우세하다고 평가가 되었다.

### 2.3 배경과 인물의 합성 기법

세 번째 기법으로는, 제3장에서 언급한 Remove.bg와 같은 상용 도구를 활용하여 배경과 인물을 합성하는 방안을 시도하였다.

그림 15는 영화 속 주인공이 아침 출근을 준비하며 화장실에서 첫 솔질하는 장면이다. 당초 실제 고흐(Van Gogh) 작품을 직접 합성하기 어려워 “고흐 풍의 그림”이라는 텍스트 제시어를 사용했으나, 이로 인해 그림의 분위기와 실제 영상 간에 부자연스러운 이질감이 발생하였다. 또한 주인공의 후면 모습의 색감이나 영화의 분위기와는 다른 3D 애니메이션과 같은 느낌의 배경이 생성되기도 하였다.

이를 개선하기 위해 Remove.bg를 활용하여 배경을 재구성하고, 실제 고흐의 작품을 삽입함으로써 영화의 복선을 더욱 자연스럽게 암시할 수 있었다. 다만, 거울에 비치는 주인공의 모습이 구현되지 않아 합성 장면의 완성도 측면에서 아쉬움이 남는다는 점은 향후 보완이 필요하다.



그림 14. Adobe Character Animator를 이용한 2D 캐릭터 애니메이션 예시<sup>[26]</sup>

Fig. 14. Example of 2D Character Animation Using Adobe Character Animator<sup>[26]</sup>



그림 15. remove.bg 상용 도구로 재구성한 주인공이 칫솔질 하는 장면 (기존 이미지/생성 이미지)

Fig. 15. Scene of Protagonist Brushing Teeth Reconstructed Using remove.bg Commercial Tool (Original Image/Generated Image)

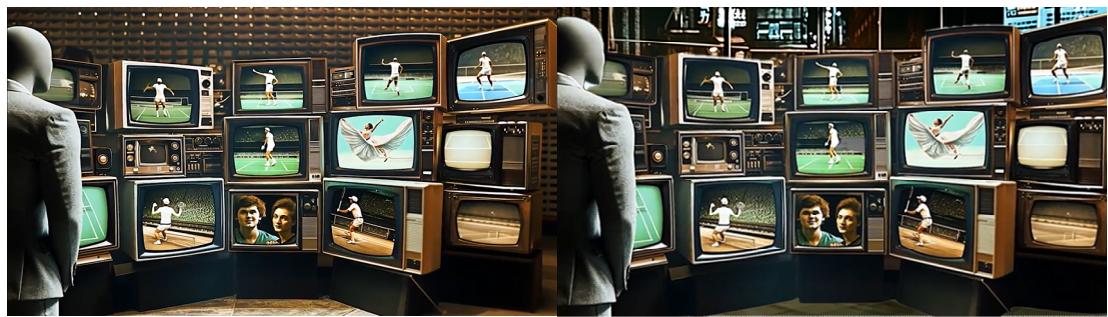


그림 16. remove.bg 상용 도구로 재구성한 주인공이 TV를 보는 장면 (기존 이미지/생성 이미지)

Fig. 16. Scene of Protagonist Watching TV Reconstructed Using remove.bg Commercial Tool (Original Image/Generated Image)

한편, 본 작품인 「The Face」는 고전 명화를 주요 소품으로 활용하였으며, 저작권에 문제가 없는 퍼블릭 도메인 작품이었기 때문에 이러한 합성 시도가 가능하였다. 소품이 중요한 역할을 하는 작품이라면 본 기법을 활용하기 수월 하지만, 저작권과 관련된 민감한 이슈가 있는 경우에는 그대로 사용하기 어려울 수 있다. 이럴 때는 이미지 생성형 AI를 활용해 새로운 이미지를 직접 제작함으로써, 저작권 문제를 회피하는 방안을 고려해야 한다.

그림 16은 영화 속 주인공이 감정 변화를 겪는 핵심 장면으로, 텔레비전에 등장하는 인물들이 각자의 꿈을 펼치는 모습을 지켜보는 장면이다. 배경은 원래 텔레비전 매장이어야 했으나, Midjourney를 통해 이미지를 생성하고 Runway의 Image to Video 모델을 사용해 합성하는 과정에서 장면이 지나치게 복잡해져, 실제 매장을 제대로 재현하지 못한 아쉬움이 있었다.

이를 보완하기 위해 remove.bg를 활용하여 기존 이미지를 분리하고, Midjourney에서 새로 제작한 배경을 합성함

으로써 이질감을 줄이고 영화의 몰입도를 높일 수 있었다.

따라서 위에서 언급한 기술을 사용해 AI로 생성된 사진의 배경에서 영상의 분위기와 맞지 않거나 어색하게 생성된 부분을 새로 AI를 통해 생성하거나 실제로 촬영된 사진을 사용해 수정, 보완을 한다면 보다 자연스러운 영상이 나올 것이다.

## V. 제작 결과 분석 및 고찰

### 1. 제작 결과 평가

본 연구에서는 AI 도구를 활용하여 단편영화를 제작하였으며, 다양한 작업에서 특정 도구들이 사용되었다. 표 1은 본 연구에서 활용된 주요 AI 도구와 그 역할을 정리한 것이다.

표 1. 단편영화 제작에서 사용된 AI 도구 개요  
 Table 1. Overview of AI Tools Used in Short Film Production

| Task                         | AI Tools Used                    |
|------------------------------|----------------------------------|
| Facial Expression Adjustment | Act-One, AI Lab Tools            |
| Motion Implementation        | DeepMotion, Flux Tools, OpenPose |
| Background Composition       | Remove.bg, Stable Diffusion      |
| Video Generation             | Runway Gen-3, Midjourney         |

본 연구에서는 표 1에서 정리된 AI 도구들을 활용하여 단편영화를 제작하고, 기존 방식과 비교 평가를 수행하였다. 특히, 기존 생성형 AI 모델(예: Runway Gen-3, Midjourney)과 최신 Pretrained Model을 비교하여, 신기술 적용이 모델 성능에 미치는 영향을 정성적으로 분석하였다. 이를 통해, AI 기반 제작 방식이 단편영화의 품질과 효율성에 어떠한 차이를 가져오는지 평가하고자 하였다.

본 연구의 평가는 논문 작성에 참여한 팀원 8명을 대상으로 진행되었다. 각 팀원은 신기술 적용 전후의 모델 성능을

비교하며, 사용성(Ease of Use), 결과의 정확성(Accuracy of Results), 출력 세부 사항(Detail in Outputs), 전반적인 평가(Overall Evaluation)를 기준으로 5점 척도로 점수를 매겼다. 이를 통해, 신기술을 적용한 모델이 사용자 만족도 및 세부 출력에서 높은 평가를 받았지만, 일부 조건에서는 기존 모델보다 사용성이 떨어지는 결과도 확인되었다.

기존 모델은 사용성이 높지만, 섬세한 감정 표현과 복잡한 장면 구성에서 한계를 드러냈으며, 결과물의 품질이 일관되지 않아 반복적인 프롬프트 수정이 필요했다. 반면, 신기술을 적용한 DeepMotion과 Flux Tools는 행동 표현과 이미지 품질 개선에서 우수한 성능을 보였고, Stable Diffusion은 배경 및 구성 편집 작업에서 높은 유용성을 나타냈다.

평가 결과, 최신 Pretrained Model을 활용한 신기술 적용 모델이 정확성과 디테일에서 기존 모델을 능가하였다. 각 모델은 고유한 강점과 한계를 가지며, DeepMotion과 Flux Tools는 동작 표현과 이미지 품질 개선에, Stable Diffusion

표 2. 신기술 적용 전후 모델 성능의 정성평가  
 Table 2. Qualitative Evaluation of Model Performance Before and After Applying New Technology

|                           | When new technology is not applied(Gen-3, Midjourney)   | When new technology is applied |   | Ease of Use | Accuracy of Results | Detail in Outputs | Overall Evaluation |
|---------------------------|---|--------------------------------|---|-------------|---------------------|-------------------|--------------------|
| Facial expression         | The user creates the prompt himself and repeatedly executes it until the desired result is obtained to derive the result. However, the quality of the result value is not constant. | AI Lab                         | By using the pretrained model, it is possible to generate high-quality images in various ways with desired elements or saturation colors of the background. | 4.5/5       | 4.7/5               | 3.2/5             | 4.0/5              |
|                           |   | Act-One                        | If you have the desired reference image, you can set various expressions as you want.   | 4.7/5       | 4.8/5               | 4.5/5             | 4.6/5              |
| Behavior                  | The user creates the prompt himself and repeatedly executes it until the desired result is obtained to derive the result. However, the quality of the result value is not constant. | Deep motion                    | The desired motion can be implemented using mannequins, and the accuracy and quality of the results are good.   | 4.6/5       | 4.8/5               | 4.5/5             | 4.6/5              |
|                           |   | Adobe character animator       | It generates images through real-time motion capture, but has limitations in some features.   | 4.0/5       | 4.0/5               | 3.5/5             | 3.0/5              |
|                           |   | Flux tools                     | It improves image quality effectively, and improves the overall detail of the image.  | 4.5/5       | 4.7/5               | 4.8/5             | 4.7/5              |
| Background and Characters | The user creates the prompt himself and repeatedly executes it until the desired result is obtained to derive the result. However, the quality of the result value is not constant. | Stable Diffusion               | It is highly flexible because certain areas and components can be additionally edited.  | 4.4/5       | 4.6/5               | 4.2/5             | 4.3/5              |
|                           |   | Remove.bg                      | It provides a function that allows a user to select and synthesize an image desired.  | 4.8/5       | 4.7/5               | 4.6/5             | 4.7/5              |

은 배경 편집 작업에 적합한 것으로 평가되었다.

## 2. 제작 결과 고찰

본 연구에서는 제안한 세 가지 영화 제작 기법을 「The Face」 제작 과정에 적용하여, 기존 생성형 AI 도구에서 나타난 한계를 개선하고자 하였다. 각 기법의 적용 결과와 개선점을 다음과 같이 요약할 수 있다.

## 3. 표정 연기 기술

기존 AI 도구는 감정을 세밀하게 표현하지 못하고, 무표정이거나 과장된 표정을 생성하는 문제가 있었다. AILab Tools를 사용해 표정 생성 과정을 단순화할 수 있었으나, 프리셋 기반 변형만 가능하여 미묘한 감정 표현에는 한계가 있었다. 반면, Act-One은 레퍼런스 영상을 활용해 자연스럽고 일관된 표정을 생성할 수 있었으나, 적합한 레퍼런스 영상을 찾는 데 시간이 소요되었다. 이러한 한계는 표정 연기 기술의 향상을 통해 일부 개선되었으나, 보다 세밀한 감정 표현을 위해 추가적인 도구와 접근법이 필요하다.

## 4. 행동 입히기 및 애니메이션 적용 기법

기존 도구로는 손·발의 디테일이 부족하거나 동작이 왜곡되는 문제가 빈번했다. DeepMotion을 활용하여 Skeleton 기반 동작 구현을 시도한 결과, 손가락 각도와 동작이 더 자연스럽게 표현되었다. Flux Canny 모델과 OpenPose를 결합함으로써 외곽선 정보와 자세 데이터를 활용해 정교한 동작 구현이 가능해졌다. Adobe Character Animator는 간단한 2D 콘텐츠 제작에 효과적이었으나, 실사형 동작 구현에는 적합하지 않았다. 행동 표현과 애니메이션 품질은 전반적으로 개선되었으나, 일부 기능은 유료화로 인해 제한적으로 사용되었다.

## 5. 배경과 인물의 합성 기법

기존 도구는 배경과 인물의 자연스러운 조화에 어려움이 있었고, 특정 스타일을 생성하는 데 있어 이질감이 발생하

였다. Remove.bg와 같은 도구를 활용해 배경을 재구성하고, Stable Diffusion과 MidJourney를 결합하여 보다 자연스러운 장면 합성을 구현하였다. 특히 고전 명화를 삽입하거나 텍스트 기반 배경 생성을 통해 독특한 연출이 가능해졌으나, 일부 장면에서는 거울 반사나 복잡한 배경의 정확성이 부족하여 완성도가 제한되었다. 배경과 인물의 합성은 기술적 보완을 통해 상당 부분 개선되었으나, 특정 제약은 여전히 존재한다.

## VI. 결 론

본 연구에서는 생성형 AI를 활용한 단편영화 제작에서 발생하는 주요 한계를 극복하기 위해 세 가지 기법(표정 연기, 동작 구현 및 애니메이션 적용, 배경 합성)을 적용하고 그 효과를 분석하였다. 연구 결과, AI 기반 영화 제작 방식은 기존 제작 방식과 비교하여 제작의 효율성을 높이고 기술적 접근성을 확장하는 가능성을 보여주었으나, 감정 표현, 동작의 자연스러움, 배경과 인물의 조화 등에서 여전히 한계가 존재함을 확인할 수 있었다.

본 연구에서 가설로 제시한 “생성형 AI가 단편영화 제작 과정에서 창의성과 생산성을 증진할 수 있다”라는 점은 부분적으로 입증되었다. AI 도구들은 제작 과정을 자동화하고 반복적인 작업을 줄이는 데 기여하였으나, 정밀한 감정 표현과 세밀한 동작 구현에서는 기존의 인간 창작자의 역할을 완전히 대체하기 어려운 것으로 나타났다. 특히 AI가 생성하는 영상 요소들은 데이터 학습에 의존하기 때문에, 학습 데이터의 품질과 범위에 따라 결과물의 일관성이 달라지는 문제가 있었다.

본 연구의 한계로는 첫째, 실험에 사용된 AI 도구가 특정 상용 도구 및 사전 학습된 모델(Pretrained Model)에 국한되어 있어, 보다 다양한 도구와의 성능 비교가 이루어지지 않았다는 점을 들 수 있다. 둘째, 일부 AI 도구들은 사용자의 숙련도에 따라 결과물의 품질이 달라질 수 있어, 동일한 도구를 활용하더라도 제작자의 개입 정도에 따라 결과가 달라지는 문제가 있었다. 셋째, 본 연구에서는 단편영화 제작을 중심으로 분석이 이루어졌으나, 장편영화나 특정 장르의 영화에서는 생성형 AI 기술이 어떻게 적용될 수 있는

지에 대한 논의가 부족하였다.

향후 연구에서는 AI 기술이 영화 제작의 보다 다양한 단계에 걸쳐 효과적으로 활용될 수 있는 방안을 모색해야 한다. 특히, 더욱 정교한 감정 표현과 자연스러운 동작 구현을 위한 AI 모델 개선, 다양한 영화 장르와 스타일에 대한 AI의 적용 가능성 평가, 그리고 AI 기술이 상업적 영화 제작에 미치는 영향을 분석하는 연구가 필요할 것이다. 또한, AI가 창작 과정에서 인간 창작자의 역할을 어떻게 보완할 수 있는지, 그리고 AI 기반 콘텐츠 제작의 윤리적·법적 쟁점에 관한 연구도 중요한 과제가 될 것이다.

이 연구는 생성형 AI를 활용한 영화 제작이 기존 제작 방식과 어떻게 조화를 이루며 발전할 수 있는지에 대한 기초 연구로서, 향후 영화 제작 과정에서 AI 기술의 활용 가능성을 보다 체계적으로 평가하는 기반이 될 것으로 기대된다.

## 참 고 문 헌 (References)

- [1] Magenta, <https://magenta.tensorflow.org/> (accessed Feb 19, 2025)
- [2] Samsung SDS, The Story on Emotional AI, <https://www.samsungsds.com/kr/insights/emotional-ai.html> (accessed Feb 19, 2025)
- [3] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Freeman, “Image-to-Image Translation with Conditional Adversarial Networks,” arXiv preprint arXiv:1611.07004, 2017, (accessed Mar. 10, 2025).  
doi: <https://doi.org/10.48550/arXiv.1611.07004>
- [4] The Face, <https://youtu.be/bY0PgdAwaxg?si=NIXsB-bcNnfTMt57> (accessed Feb 24, 2025)
- [5] W. Benjamin, The Work of Art in the Age of Its Mechanical Reproducibility, in Illuminations, H. Arendt, Ed., New York: Schocken Books, pp.217-251, 1964.
- [6] L. Manovich, “Artificial Aesthetics: Separate and Reassemble: Generative AI Through The Lens of Art and Media Histories,” ResearchGate, 2023, (accessed Mar. 10, 2025).  
doi: <https://doi.org/10.11606/issn.1982-8160.v18i2p7-18>
- [7] G. Simondon, On the Mode of Existence of Technical Objects, C. Malaspina and J. Rogove, Trans., Univocal Publishing, 2017.
- [8] Elevenlabs, <https://elevenlabs.io/> (accessed Jan. 7, 2025)
- [9] Adobe Premiere Pro, <https://www.adobe.com/kr/products/premiere.html> (accessed Jan. 11, 2025)
- [10] Midjourney, <https://www.midjourney.com/imagine> (accessed Jan. 11, 2025)
- [11] Goodfellow, Ian, et al. “Generative adversarial nets.” Advances in neural information processing systems, 2014.  
doi: <https://doi.org/10.48550/arXiv.1406.2661>
- [12] Qiao, Fengchun, et al. “Geometry-Contrastive GAN for Facial Expression Transfer.” arXiv preprint, 2018.  
doi: <https://doi.org/10.48550/arXiv.1802.01822>
- [13] Karras, Tero, et al. “A style-based generator architecture for generative adversarial networks.” IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.  
doi: <https://doi.org/10.48550/arXiv.1812.04948>
- [14] AILab Tools, <https://www.ailabtools.com/> (accessed Jan. 7, 2025)
- [15] Hunyuan Video, <https://aivideo.hunyuan.tencent.com/> (accessed Jan. 7, 2025)
- [16] flux tool, [https://flux-ai.io/?gad\\_source=1&gclid=Cj0KCQiAq-u9BhCjARIIsANLj-s1L7zprdk8coSkCVFRpQW03d9wjdZZIV1V7AQ5JV\\_LfiMnbBAygRtXwaAsuVEALw\\_weB](https://flux-ai.io/?gad_source=1&gclid=Cj0KCQiAq-u9BhCjARIIsANLj-s1L7zprdk8coSkCVFRpQW03d9wjdZZIV1V7AQ5JV_LfiMnbBAygRtXwaAsuVEALw_weB) (accessed Jan. 7, 2025)
- [17] Cao, Zhe, et al. “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, 2021, pp. 172-186.  
doi: <https://doi.org/10.1109/TPAMI.2019.2929257>
- [18] Hunyuan Video Github, <https://github.com/Tencent/HunyuanVideo> (accessed Jan. 7, 2025)
- [19] Blackforestlabs, <https://blackforestlabs.ai/flux-1-tools/> (accessed Jan. 7, 2025)
- [20] “Flux Dev OpenPose ControlNet.” Hugging face, [https://huggingface.co/raulc0399/flux\\_dev\\_openpose\\_controlnet](https://huggingface.co/raulc0399/flux_dev_openpose_controlnet) (accessed Jan. 13, 2025)
- [21] Stable Diffusion, <https://stablediffusionweb.com/ko> (accessed Jan 11, 2025)
- [22] Remove.bg, <https://www.remove.bg/ko/upload> (accessed Jan.11, 2025)
- [23] Deepmotion, [https://www.deepmotion.com/?gad\\_source=1&gclid=CjwKCAiAp4O8BhAkEiwAqv2UqAf-U9UL2Rr0veeccc9Pj3bFf0KT840Rx1XUnhrKR5WZm7JYTkT7xoCmnoQAvD\\_BwE](https://www.deepmotion.com/?gad_source=1&gclid=CjwKCAiAp4O8BhAkEiwAqv2UqAf-U9UL2Rr0veeccc9Pj3bFf0KT840Rx1XUnhrKR5WZm7JYTkT7xoCmnoQAvD_BwE) (accessed Jan. 10, 2025)
- [24] Z. Wang, et al., “ComfyBench: Benchmarking LLM-based Agents in ComfyUI for Autonomously Designing Collaborative AI Systems,” arXiv preprint, arXiv:2409.01392, 2024. (accessed Jan. 13, 2025)  
doi: <https://doi.org/10.48550/arXiv.2409.01392>
- [25] Openpose Github, <https://github.com/CMU-Perceptual-Computing-Lab/openpose?tab=readme-ov-file> (accessed Jan. 7, 2025)
- [26] Adobe Character Animator, <https://www.adobe.com/kr/products/character-animator.html> (accessed Jan. 11, 2025)

---

### 저자소개

---



김채연

- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0003-4707-6256>
- 주관심분야 : 생성형 AI 영화 제작, 영상콘텐츠 제작, 방송기술



주하은

- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0004-0186-9362>
- 주관심분야 : Computer Vision, Generative AI, 방송기술, 방송통신



황진성

- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0008-4304-435X>
- 주관심분야 : 영상콘텐츠 제작, 인공지능



유성권

- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0005-2674-9903>
- 주관심분야 : 영상처리, 인공지능



명윤호

- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0004-7858-2474>
- 주관심분야 : Deep Learning, Generative Model

---

### 저자소개

---

#### 임영래



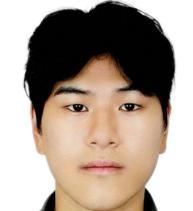
- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0005-7040-8276>
- 주관심분야 : 영상처리, 인공지능

#### 인예원



- 서울과학기술대학교 스마트ICT융합공학과 학부생
- ORCID : <https://orcid.org/0009-0004-6727-6324>
- 주관심분야 : 인공지능 대규모 언어 모델(LLM)

#### 김다루



- 서울과학기술대학교 전자T미디어공학과 학부생
- ORCID : <https://orcid.org/0009-0007-8441-3193>
- 주관심분야 : 영상처리, 인공지능 미디어

#### 이승민



- 서울과학기술대학교 문예창작학과 학부생
- ORCID : <https://orcid.org/0009-0002-6457-0654>
- 주관심분야 : 영화 시나리오, 미디어와 영상문학

#### 박구만



- 1984년 2월 : 한국항공대학교 전자공학과
- 1986년 2월 : 연세대학교 전자공학과 공학석사
- 1991년 2월 : 연세대학교 전자공학과 공학박사
- 1991년 3월 ~ 1996년 9월 : 삼성전자 신호처리연구소 선임연구원
- 2006년 1월 ~ 2007년 8월 : Georgia Institute of Technology Dept. of Electrical and Computer Engineering, Visiting Scholar
- 2016년 1월 ~ 2017년 12월 : 서울과학기술대학교 나노IT디자인융합대학원
- 1999년 8월 ~ 현재 : 서울과학기술대학교 스마트ICT융합공학과 교수
- ORCID : <https://orcid.org/0000-0002-7055-5568>
- 주관심분야 : 컴퓨터비전, 지능형실감미디어

---

저자소개

---

이영화



- 2001년 8월 ~ 2021년 12월 : 현대홈쇼핑 방송그래픽팀 NLE 감독
- 2024년 2월 : 서울과학기술대학교 전자T미디어공학과 박사수료
- 2022년 1월 ~ 현재 : 서울과학기술대학교 신학협력단 IoT융복합연구소 책임연구원
- ORCID : <https://orcid.org/0000-0002-2427-7667>
- 주관심분야 : Computer Vision & Graphics, Generative AI, Video-Language pre-training

양지희



- 2012년 2월 : 강릉원주대학교 멀티미디어공학과 공학사
- 2015년 2월 : 서울과학기술대학교 미디어IT공학과 석사
- 2015년 3월 ~ 2020년 8월 : 서울과학기술대학교 나노IT디자인융합대학원 정보통신미디어공학과 박사
- 2020년 3월 ~ 현재 : 서울과학기술대학교 IoT융복합기술연구소 선임연구원
- ORCID : <https://orcid.org/0000-0003-2776-5487>
- 주관심분야 : 컴퓨터 비전, 실감미디어

최영희



- 서울과학기술대학교 문예창작학과 조교수
- ORCID : <https://orcid.org/0009-0004-2286-5029>
- 주관심분야 : 생성형 AI 영화 제작, 영화 시나리오, 동아시아 영화 및 TV 드라마, 미디어와 영상문학