

<http://dx.doi.org/10.17703/JCCT.2024.10.3.723>

JCCT 2024-5-81

## 인공지능기반의 자동 창작 영상에 관한 논구

### A Discussion on AI-based Automated Picture Creations

김정희\*, 윤준성\*\*

Junghoe Kim\*, Joonsung Yoon\*\*

**요약** 본 연구는 자동 창작 영상에 관한 개념과 이해의 변화를 추적하기 위하여, 기존의 영상분야를 대표하는 사진과 영화의 창작 방식을 인공지능 기반의 영상 창작 방법과 '자동'이라는 측면에서 유비적으로 탐구하고, 새로운 자동영상창작에 관한 이해와 가능성을 논구 한다. 사진과 영화의 발명 시기에는, 회화를 비롯한 전통예술 장르와 비견하여, 이들에게 '자동 창작'이라는 영역을 설정한 바 있다. 최근 인공지능이 영상 제작에도 적용되면서 '자동 창작'에 관한 개념도 확장되어, 기존의 문학, 미술, 사진, 영화의 경계를 자유롭게 넘나드는 실험적 창작이 활발하다. 인공지능의 자동 창작은 머신 러닝과 딥 러닝과 같은 기술을 활용하여, 인공지능이 독립적으로 창작 과정을 수행하게 한다. 인공지능을 활용한 자동 창작은 효율성을 크게 향상시킬 수 있지만, 동시에 예술창작과정의 주관적인 판단의 문제에 직면하게 된다. 자동 창작으로 몰렸던 사진과 영화의 창작과정과 인공지능의 창작과정을 비교하면, 자동 창작 영상의 입면은 작가의 문제로 귀착될 수 있다. 그러나 다원적이 예술분야가 소통할 수 있는 기회로 작동된다면, 인공지능기반의 자동 창작 영상은 장르를 넘나드는 방법론이 될 수 있을 것이다.

**주요어** : 인공지능, 자동 창작, 젠-2, 사진, 영화

**Abstract** In order to trace the changes in the concept and understanding of automatically generated images, this study analogously explores the creative methods of photography and cinema, which represent the existing image fields, in terms of AI-based image creation methods and 'automaticity', and discusses the understanding and possibilities of new automatic image creation. At the time of the invention of photography and cinema, the field of 'automatic creation' was established for them in comparison to traditional art genres such as painting. Recently, as AI has been applied to video production, the concept of 'automatic creation' has been expanded, and experimental creations that freely cross the boundaries of literature, art, photography, and film are active. By utilizing technologies such as machine learning and deep learning, AI automated creation allows AI to perform the creative process independently. Automated creation using AI can greatly improve efficiency, but it also risks compromising the personal and subjective nature of art. The problem stems from the fact that AI cannot completely replace human creativity.

**Key words** : AI, automatic creation, Gen-2, Photography, Cinema

\*정희원, 숭실대학교 대학원 미디어학과 박사수료 (제1저자)

\*\*정희원, 숭실대학교 대학원 미디어학과 교수 (교신저자)

접수일: 2024년 3월 4일, 수정완료일: 2024년 4월 10일

게재확정일: 2024년 4월 20일

Received: March 4, 2024 / Revised: April 10, 2024

Accepted: April 20, 2024

\*\*Corresponding Author: jsy@ssu.ac.kr

Dept. of Digital Media, Graduate School of Soongsil Univ, Korea

## I. 서론

15세기 이후부터 발명과 사회의 조직화는 상호 긴밀하게 영향을 미쳤다. 기계, 제분기, 시계, 총, 인간 같은 자동인형 등의 수와 종류가 증가하면서, 인간에게 유용한 기계적 속성들이 제시되었고, 더욱 미묘하고 복잡한 유기적 사실을 기계적 비유로 해석하려는 경향도 늘어났다 [1]. 이러한 경향은 인간과 기계, 자연과 인공 등의 유비와 혼합을 촉진했으며, 자동화라는 의미를 통해 인간의 개입이 축소되는 상황에 대한 유용성이 논의되고 있다.

19세기와 20세기 초 사진과 영화의 발명 시기에는 예술의 ‘자동 창작’ 영역을 새롭게 정의하며, 이는 회화와 다른 전통 예술 형태와의 차별화를 명확히 하였다. 현대 예술은 지속적인 기술적 발전의 영향 아래에서 빠르게 변화하고 있으며, 이러한 변화의 중심에는 ‘자동 창작’이라는 개념이 자리 잡고 있다. 사진과 영화의 등장은 이 변화의 초기 단계를 드러내는데, 이전까지 예술은 주로 회화나 조각과 같은 수작업에 의존하였기 때문이다. 그러나 사진과 영화의 도입으로 기계와 기술이 창작 과정에서 큰 역할을 하게 되면서 예술의 방식과 그 의미에 근본적인 변화가 시작되었다.

사진과 영화의 등장은 시각 예술 분야에 중요한 영향을 주었다. 이와 유사한 시각적 표현은 주로 그림, 조각, 건축 같은 전통 예술 형태에 의존했다. 사진과 영화의 발명은 현실을 일부 자동으로 포착하는 새로운 인식 체계를 제시하고자 한다. 사진과 영화의 자동화는 혁명적이었으며, 과거에는 이를 수 없는 정확한 재현으로 시각적 장면을 복제하였다.

카메라는 기계적, 화학적 과정을 통해 시간을 순간 포착하였으며, 그림에서 볼 수 있는 주관적 해석을 효과적으로 표현하였다. 예술의 본질, 예술가의 역할, 표현의 진정성에 대한 논의와 토론이 이루어졌다. 전통주의자들은 이러한 새로운 매체를 회의적으로 바라보았으며, 그것을 손으로 그린 예술보다 ‘진정성’이 떨어진다고 생각했다. 그러나 다른 예술가들은 그들이 제시한 가능성을 받아들였으며, 새로운 예술 운동과 장르의 확장으로 발전하였다.

이러한 변화는 사진과 영화의 매체였던 필름이 디지털 화면으로 전환되는 과정을 거치면서 더욱 확장되는 영상의 영역을 보여주었다. 최근 인공지능의 등장과 함께 영상분야에서의 창작과정의 변화되는 속도는 더욱 빨라졌

다. 인공지능은 ‘자동 창작’의 개념을 사진과 영화의 자동화를 넘어서고 있다. <그림 1>과 같이 이미지나 영상을 기록하는 것을 넘어, 인공지능은 문학적, 미술적 창작과 표현, 그리고 영상 제작과 같은 다양한 예술 형태에서 중요한 역할을 하게 된다.

창작자로서 인간이 수행해 온 전형적이고 전통적인 역할을 인공지능이 대신할 수 있는 상황이 도래했다. 이러한 예술의 영역은 전통적인 경계들이 허물어지며, 다양한 예술 형태와 기술이 서로 융합되는 새로운 패러다임을 맞이할 수 있는 가능성을 가지게 되었다. 예술가들에게 전례 없는 창작의 다원적인 가능성을 부여하며 예술의 본질과 그 방향성에 대한 깊은 재성찰을 필요로 한다.



그림 1. 추상 0002, 젠-2, 텍스트 투 이미지, 2023 © 김정호.  
Figure 1. Abstract 0002, Gen-2, Text to Image, 2023  
© JungHoe Kim.

## II. 사진과 영화라는 자동 창작

1839년 프랑스 과학 아카데미는 은판 사진 기술의 발명을 공표했다[2]. 사진이라는 표현매체의 특징을 ‘자동’의 측면으로 밀어내는 경향이 보인다. 이는 기존의 인물화, 풍경화, 정물화 등의 회화와 대비되는 창작방식에 연유한다. 수동 창작(Manual Creation)은 예술가의 직접적인 물리적 참여와 주관적인 표현이 중심이 되는 예술 창작 방식이다. 이러한 방식은 예술가의 감정, 경험, 세계관이 작품에 깊숙이 반영되었으며, 그 과정 자체도 중요한 예술적 가치를 지니고 있다. 예술가는 다양한 재료와 도구를 사용하여 자신만의 독특한 스타일과 표현 방식으로 작품을 창출한다. 예를 들어, 화가는 점, 선, 면, 양감, 부피, 질감, 명도, 색상, 형태, 구성 등과 같은 요소를 조합

하여 이미지를 창조하며, 화가의 관찰, 상상력, 그리고 기술이 결합된 결과물이다. 또한, 조각가는 다양한 재료를 깎고, 모양을 내며, 구조를 만든다. 이처럼 수동 창작은 예술가의 내면과 외부 세계 사이의 상호작용을 통해 이루어지며 각각의 작품들은 개인적이며 독특한 작품 세계를 만든다. 이는 예술가와 관람자 간의 깊은 소통과 이해를 가능하게 하며, 다양한 문화와 예술적 가치를 세상에 전달할 수 있다. 수동 창작은 예술가의 개인적인 감정과 생각, 그리고 예술 세계를 표현하는 데 중요한 역할을 하며, 작품마다 독특한 가치와 의미를 만들어낸다. 이러한 측면에서 볼 때에, 사진과 영화는 당시의 전형적인 창작 방법과는 상이한 방식을 차용하며, 자동 창작과 같은 분위기를 자아내었다.

적절한 노광시간(Exposure time)을 결정해야 하는 사진기와 영상카메라는 초기에는 경험과 실수기반의 시도들로 이루어졌으나, 노광을 측정하는 기기를 통한 측정이 가능해졌고, 거리측정도 실측에서 렌즈와 사진기 기반의 측정 방법으로 일종의 간편화가 진행되었다. 20세기 중반이 지나면서 적절한 노광에 관한 문제는 자동화로 진전되어, 사진기에는 완전자동(Program Mode)으로 노광시간을 결정해 주거나, 조리개 우선(Aperture Mode) 혹은 셔터 우선(Shutter Mode)의 자동화 방법이 도입되었다. 물론 완전 수동(Manual Mode)으로 노광량을 설정하는 기능도 여전히 함께 존재하였으며, 이는 완벽한 사진가의 기법으로 여겨지기도 하였다. 자그마한 컴퓨터 칩의 도입은 완전자동 기능을 풍성하게 만드는 역할을 수행했고, 소위 자동사진기는 일반인이 자유롭게, 그러나 실수 없이 사진을 촬영할 수 있는 편의를 제공하였다. 영상카메라도 이와 유사한 발전을 하였으나, 셔터속도가 한정되어 있어 인공조명을 통한 노광시간 충족을 지향하곤 했다. 그러나, 필름기반의 사진기와 영상카메라는 필름의 감광도와 색온도에 의존해야 하였고, 감광도는 해상도와도 연관되어 있어, 작가와 감독의 결정이 절대적이었다.

아마도, 디지털사진(Digital Photography)이 그러한 자동화를 촉진했다는 논의도 간혹 있으나, 필름기반의 사진과 영화에서 발견되는 자동화개념을 넘어서지는 않는다. 디지털사진은 디지털사진기의 사용과 얻어진 이미지의 편집까지 포함하는 영역이며, 필름 현상과 인화를 거치지 않고 즉각적으로 영상의 확인하고, 필요에 따라 변형 및 보정할 수 있는 편의성이 강조된다. 디지털 파일로 생성된 영상은 사진이건, 동영상이건 파일로 저장, 복

사, 배포가 가능해졌으며, 이는 경제적, 산업적, 문화적 측면에서 커다란 변화를 가져왔다. 또한 디지털 형태의 영상은 컴퓨터 기반 영상 기술들과의 상호 발전을 가능하게 만들었다. 즉, 영상이 창작되는 방향이 쌍방향으로 전환되었다는 것이다. 실제 대상이 없어도 생산할 수 있는 컴퓨터 그래픽스의 발전은 그저 단순한 한 예일 뿐이다. 자동화라는 이름으로 용인되는 사진, 영화, 디지털영상의 창작은 여전히 수동적 측면에 기대고 있었고, 창작자의 역할은 편의성을 극복하면서도 여전히 인간의 영역에 머물고 있었다.

### III. 인공지능기반의 자동 창작

최근의 인공지능은 영상 제작 통합과 함께 시각 예술 분야에서 또 다른 변혁적인 단계를 목격하게 한다. 이 통합은 자동 창작의 경계를 재설정 하고 있다. 전통적인 사진과 영화에서 볼 수 있는 수동적 자동화와는 달리, 인공지능은 기계의 포착뿐만 아니라 해석, 설계, 심지어 혁신까지도 수행하는 활동적인 형태의 자동화를 도입한다. 이로 인해 문학, 예술, 사진, 영화의 요소가 때때로 결합되어 복합 형태의 예술적 표현이 나타나게 되었다. 융합과 간학제적 연구들이 추구해온 방향은 이와 맥락을 같이 하는 것으로 보인다.

인공지능이 영상 제작에서 수행하는 역할은 다양하다. 편집부터 색 보정에 이르기까지의 과정을 간소화하여 작업을 더 효율적으로 만든다. 반면에, 인공지능은 스스로 창작자가 될 가능성을 보여주고 있다. 예컨대, 생성 알고리즘(Generative algorithms)은 고유하고 전통적인 미학의 경계를 인간에게서 밀어내는 시각적 요소를 생성할 수 있다. 이것은 사진의 출현 시기에 볼 수 있었던 논쟁과 유사한 논쟁을 불러일으켰다.

인공지능 주도 창작의 시대에 인간 예술가의 역할은 여전히 중요하다. 인공지능은 창작 과정에서 강력한 도구로 활용될 수 있지만, 예술의 본질은 인간의 감성과 창의력에서 비롯된다. 인간 예술가는 인공지능이 생성한 결과물을 바탕으로 자신만의 해석과 비전을 더해 작품을 완성하는 역할을 수행한다. 또한 인공지능 기술의 발전 방향을 설정하고, 윤리적 기준을 설정하는데 인간 예술가의 통찰력이 필요하다.

인공지능으로 생성된 예술이 '진정한' 예술인지에 대해서는 현재 명확한 합의가 이루어지지 않았다. 하지만

예술의 가치는 단순히 창작 주체가 인간인지 기계인지에 따라 결정되는 것이 아니라, 작품 자체가 지닌 미학적 가치와 메시지, 그리고 감상자에게 주는 영감과 공감에 의해 평가되어야 합니다. 인공지능 생성 예술도 이러한 기준에서 충분히 '진정한' 예술로 인정받을 수 있다.

새로운 맥락에서 창의성은 인간과 인공지능의 협업을 통해 재 정의될 수 있다. 인공지능은 방대한 데이터를 바탕으로 새로운 아이디어와 표현 방식을 제시할 수 있으며, 인간은 이를 자신만의 감성으로 해석하고 발전시킬 수 있다. 이러한 과정에서 인간과 인공지능은 서로의 강점을 활용하여 기존에 없었던 창의적 결과물을 만들어 낸다. 창의성은 인간과 인공지능의 상호작용과 협력을 통해 확장되고 진화하는 개념으로 이해될 수 있다.

### 1. 자동 창작 영상의 발달과 변화

영상 창작에서의 인공지능 응용은 단순한 기술적 진보가 아니다. 그것은 문화적이고 철학적인 변화가 될 수 있다. 그것은 우리의 저작권, 창의성, 예술의 본질에 대한 관념에 도전한다. AI가 계속 발전하고 예술에 통합되는 것이 더 깊어짐에 따라, 이러한 질문은 자동화 시대의 시각 예술의 미래와 창의성의 본질에 대한 논의의 중심이 될 것이다.

반면에 현재 우리가 일컬을 수 있는 자동 창작 (Automatic Creation)은 기술과 알고리즘의 활용이 특징으로, 예술가의 직접적인 개입이 최소화되고, 기술적인 파라미터와 데이터가 중심이 된다. 이 방식은 기존의 예술 형식과 창작 과정을 확장하며, 예술과 기술의 경계를 모호하게 만든다. 예를 들어, 인공지능은 대규모의 데이터를 학습하여 새로운 이미지나 텍스트를 생성할 수 있다. 또한, 컴퓨터 그래픽스는 알고리즘과 규칙에 따라 이미지나 패턴을 자동으로 생성한다. 이러한 기술적인 접근은 예술의 새로운 영역을 개척하며, 전통적인 예술 형식에 새로운 가능성과 차원을 추가한다.

자동 창작은 예술가가 기술과 상호 작용하면서, 새로운 창작의 영역을 탐색하게 한다. 이는 예술가가 기술을 활용하여 새로운 형태의 예술 작품을 탐구하고, 기존의 예술 형태와는 다른 새로운 표현 방식을 찾아가게 된다. 자동 창작은 예술의 정의와 가치, 그리고 창작의 의미에 대해 새로운 질문을 던지며, 현대 예술의 다양한 흐름과 실험적인 작업을 가능하게 한다. 이는 미래의 예술 세계에서 새로운 장르와 형식, 그리고 예술가와 관람자의 새

로운 관계를 만들어낼 수 있다.

최근 인공지능기반 기술들의 발달이 급속히 진전하면서, 영상의 자동 창작은 인공지능이 인간의 입력 없이 작품을 창출하는 과정을 의미하게 되었다. 초기의 자동 창작은 주로 기존 영상의 재편집이나 간단한 효과 추가와 같은 수준이었으나, 현재는 인공지능 기반의 기술들이 도입되면서 그 범위와 깊이가 크게 확장되었다.

초기 자동 창작 기술은 대체로 기존의 영상 클립들을 특정 알고리즘에 따라 재배열하거나, 간단한 효과를 추가하는 정도였다. 이는 창의적 측면에서는 많은 제약이 있었으며, 대부분의 창작 과정이 사람에 의해 이루어져야 했다. 그러나 이 특정 알고리즘에 관한 연구가 주기적으로 제시되고 활성화되면서, 창작이라는 영역으로 점차 번져 나아갔다. 특히, 기계 학습 알고리즘을 사용하여 영상 데이터를 분석하고, 이를 기반으로 새로운 영상을 생성하는 연구가 활발히 진행되기 시작했다. 나아가 GAN (Generative Adversarial Network)과 같은 기술은 컴퓨터 혼자서 현실적인 이미지를 창조하게 하는 혼령 방법이다[3]. 기존의 영상 데이터를 학습하여 새로운 영상을 생성할 수 있는 능력을 보여주었다.

예를 들어, 인공지능은 다양한 스타일의 영상을 학습하여 그 스타일을 새로운 영상에 적용하는 스타일 전송 (Style transfer) 기술을 구현할 수 있다. 또한, 스토리텔링, 캐릭터 생성, 배경 디자인 등 다양한 창작 활동에서도 활용되며, 전체적인 영상 제작 과정에 혁신을 가져오고 있다. 현재의 자동 창작 기술은 단순한 작업 자동화를 넘어서, 창작의 주체로서의 역할을 수행하기 시작하였다. 이는 영상 제작의 접근성을 높이고, 새로운 형식의 작품과 표현 방식을 탐구하는 여지를 제공하며, 미래의 영상 예술에 대한 새로운 비전을 제시한다.

### 2. 인공지능기반 자동 창작 영상 분석

인공지능의 자동 창작은 영상 분야에서 혁신적인 변화를 가져오고 있다. 인공지능이 독립적으로 창작 과정을 수행하며, 딥 러닝 (Deep Learning)과 머신 러닝 (Machine Learning) 접근법이 빠르게 확산되어 새로운 형태의 작품을 생성한다는 점에서 이러한 변화가 명확하다.[4] 인공지능은 기존의 창작 방식과 달리, 시공간적인 제약을 크게 줄이고, 무한한 창작의 가능성을 열 수 있다.

예를 들어, OpenAI의 ChatGPT가 소개한 기술들은

문장 생성을 통해 새로운 문학 작품과 이미지를 창출할 수 있으며, 이는 단순한 문장 생성과 이미지 변환을 넘어서, 특정 주제나 스타일에 맞는 창작물을 만들어낼 수 있다.

OpenAI의 DALL-E 2와 DALL-E 3은 텍스트를 기반으로 이미지 생성을 수행하는 인공지능 모델의 연속된 업그레이드 버전이다. DALL-E 2는 주어진 단어와 문구에 대한 충실도와 이미지 품질 개선을 중점적으로 목표로 개발되었다. <그림 2>는 DALL-E 2를 이용해 만들어진 작품이며, 추상적인 표현과 다채로운 색채를 보여 준다.

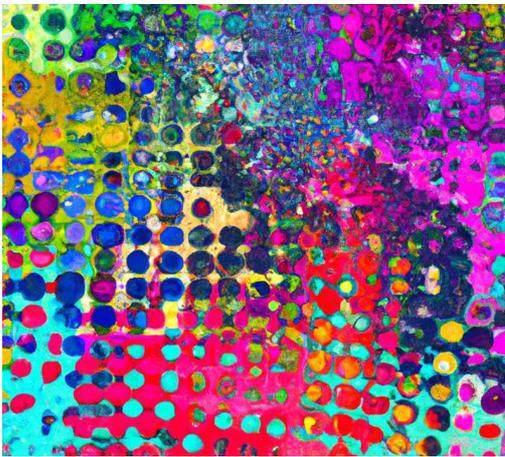


그림 2. 추상 0015, 달-이 2, 텍스트 투 이미지, 2023 © 정희.

Figure 2. Abstract 0015, Dall-E 2, Text to Image, 2023 © JungHoe Kim.

DALL-E 3는 이전 버전의 기능을 더욱 확장하고 향상시켜, 문맥의 이해능력을 개선하고, ChatGPT와의 원활한 통합을 제공하며, 강력한 보안 조치를 도입하였고, 사용자가 제공하는 미묘함과 세부 사항을 더 잘 이해할 수 있게 하였다. 또한, DALL-E 3는 ChatGPT와의 통합을 통해 예술 창작 과정에 글과 이미지가 연동할 수 있게 만들어, 더 넓은 대중에게 접근할 수 있게 하였다. 즉, 글을 통해 시각적 산물을 산출하는 방법이 선호된 상황을 만들었으며, 이는 더 나아가 영상에서 문학작품을 산출할 가능성을 시사한다. 2023년 9월 25일, OpenAI는 “ChatGPT가 이제는 보고, 듣고, 말할 수 있습니다”라고 발표하였고, 이는 기존의 예술 장르를 넘나들 수 있다는 중요성을 갖는다.

텍스트 투 이미지(Text to Image) 생성은 주어진 텍스트 프롬프트, 즉 사용자가 제시하는 문장으로부터 최신의 이미지들을 활용하여 작품을 생성하는데 초점을 맞춘다[5]. 이러한 경향은 영상을 활용하는 인공지능 회사들에게도 유사한 것이다. Runway의 Gen-1과 같은 기술은 실제 영상으로부터 애니메이션을 생성하며, Gen-2는 문장으로부터 영상을 생성한다. 특히, Gen-2는 인공지능을 이용해 텍스트, 이미지 또는 동영상을 기반으로 새로운 영상을 생성하는 다기능 모델이다. 이 도구는 사용자가 텍스트 프롬프트를 사용하여 원하는 스타일로 비디오를 만들 수 있게 해주며, 설정을 조정하고 생성물을 미세 조정할 수 있고, Runway라는 클라우드 기반 영상 편집기를 통해 영상을 다운로드하거나 저장할 수 있다. 또한 Gen-2는 고해상도의 실감나고 일관된 비디오를 생성할 수 있어, 시장에서 처음으로 공개된 텍스트 투 비디오(Text to Video) 모델이라고 회사 측은 주장하고 있다. Gen-2의 기능은 다양한 모드로 나뉘며, 각 모드는 다른 방식의 영상 생성을 지원한다. 예를 들어, ‘텍스트 투 비디오 모드’는 텍스트 프롬프트만을 사용하여 어떤 스타일의 비디오라도 생성할 수 있게 해준다.



그림 3. 추상 0001, 젠-2, 텍스트+이미지 투 비디오, 2023 © 김정희 .

Figure 3. Abstract 0001, Gen-2, Text+Image to Video, 2023 © JungHoe Kim.

또한 ‘텍스트 + 이미지 투 비디오(Text + Image to Video)모드’ <그림 3>은 주로 이미지와 텍스트 프롬프트를 사용하여 비디오를 생성하며, ‘이미지 투 비디오’(Image to Video) 모드는 단순히 주도 이미지만을 사용하여 비디오를 생성한다. ‘스타일라이제이션’(Stylization) 모드는 어떤 이미지나 프롬프트의 스타

일을 비디오의 모든 프레임에 전달할 수 있으며, ‘스토리보드’(Storyboard) 모드는 모의 이미지를 완전히 스타일화하고 애니메이션 렌더링으로 변환한다. 이 외에도 ‘마스크’(Mask), ‘렌더’(Render), 그리고 ‘커스터마이제이션’(Customization) 모드와 같은 다양한 기능을 통해 사용자는 Gen-2의 전체 기능을 활용하여 높은 해상도의 결과물을 얻을 수 있다. 그러나 결과물은 때때로 일관되지 않고 이질감이 느껴질 수 있다.

그 외 스테이블 디퓨전(Stable Diffusion)도 스테빌리티 AI (Stability AI)가 루트비히-막시밀리안 뮌헨 대학교(LMU Munich)의 CompVis 그룹과 런웨이의 협력 하에, 2022년 8월에 출시한 ‘텍스트 투 이미지’ 확산 모델 중 하나이다[5]. 주로 텍스트 설명에 기반하여 상세한 이미지를 생성하는 데 사용된다. 스테이블 디퓨전은 기본 생성 모델로 잠재 확산(Latent Diffusion)을 사용한다는 점에서 앞서 설명한 텍스트 투 이미지 모델 구조들과 차이가 있다[5]. 잠재 확산 모델(LDM)은 확산 모델을 오토 인코더로 감싸서 확산 과정이 이미지 자체가 아니라 이미지의 잠재 공간 표현에서 작동한다[5]. 이 획기적인 기술은 전체 이미지에서 작동하는 U-Net 모델보다 좁은 제거 U-Net 모델을 상대적으로 가볍게 유지할 수 있음을 의미한다. 오토 인코더에서 이미지 디테일을 잠재 공간으로 인코딩하고 잠재 공간을 고해상도 이미지로 다시 디코딩하는 무거운 작업을 처리하므로 확산 모델은 개념적인 잠재 공간에서만 수행된다. 이를 통해 훈련 과정의 속도와 성능이 크게 향상된다[5].



그림 4. 추상 0005, 스테이블 디퓨전, 텍스트 투 이미지, 2023 © 김정희.  
Figure 4. Abstract 0005, Stable Diffusion, Text to Image, 2023 © JungHoe Kim.

스테이블 디퓨전은 텍스트를 이미지로 변환하는 능력을 가지고 있어, 사용자가 텍스트를 입력하여 설명하면 모델이 해당 텍스트에 따라 상세한 이미지를 생성하거나 이미지를 수정하게 된다. 이 모델은 텍스트 프롬프트를 기반으로 사실적이고 설명적인 이미지를 생성하는 데 사용한다. 또한 스테이블 디퓨전은 이미지 구성, 얼굴 생성, 텍스트 생성 기능을 강화하고 이미지를 다양한 방법으로 수정할 수 있는 기능을 제공한다.<그림 4> 이 모델은 또한 저해상도 또는 저디테일 이미지의 세부 사항을 채우는 기능을 제공하며, 이러한 기능은 이미지 품질을 향상시키고 더 많은 디테일을 제공한다. 이 모델은 디퓨전 모델로 분류되며, 이는 매우 잡음이 많은 이미지에서 점차적으로 잡음을 제거하여 이미지를 생성하도록 학습되는 모델이다. 이 과정은 “역 디퓨전(reverse diffusion)”이라고 불리며, 이는 수학적 기반을 가지고 있다.

이러한 기술의 발전은 문학과 영상 예술이 상호 변환되는 새로운 창작의 영역, 즉 “Crossing-over”를 가능하게 한다. 예를 들어, 인공지능이 작성한 문장이나 이야기를 기반으로 영상이 생성되거나, 반대로 영상의 내용을 분석하여 문장이나 이야기를 생성하는 것이 가능하다. 이는 창작의 장르 경계를 무너뜨리며, 다양한 예술 분야가 상호 작용하고 결합하는 새로운 예술 형태를 만들어낸다<그림 5>. 아마도 인공지능 기반의 자동 창작 영상이 갖는 시사점은 융합, 간학제 등으로 표현되는 다양한 분과적 학문들의 현대적 추구처럼, 기존의 예술 장르와 분과를 일종의 소통과 융합으로 유도할 가능성을 가진다.



그림 5. 추상 0020, 젠-2, 텍스트 투 이미지, 2023 © 김정희.  
Figure 5. Abstract 0020, Gen-2, Text to Image, 2023 JungHoe Kim.

### 3. 인공지능기반 자동 창작 영상의 문제점과 한계

인공 지능을 활용한 자동 창작은 효율성을 크게 향상시킬 수 있으며, 동시에 예술의 개인적이고 자유로운 해석을 훼손할 위험이 있다. 인공 지능이 인간의 창의성을 완전히 대체할 수 없다는 한계를 보여준다. 창의성의 한계는 인공 지능이 데이터와 알고리즘에 의존하여 작동하기 때문에 발생한다. 예술가의 작품에 나타나는 개인적인 경험과 감정을 온전히 표현하기는 어렵다. 또한, 인공 지능은 수집된 데이터와 알고리즘의 범위 내에서만 작동하므로, 완전히 새로운 아이디어나 스타일을 창출하는 것에 한계가 있다.

인공 지능으로 생성한 작품의 저작권 귀속 문제와 원본 데이터의 저작권 취급 문제도 복잡하게 얽혀 있다. 인공지능은 기존 데이터를 학습 기반으로 사용하므로, 그 결과물이 다른 창작 작품과 유사해질 가능성이 있다.

인공 지능의 창작 능력은 제공된 학습 데이터의 품질과 양에 크게 의존하며, 고품질의 AI 모델을 학습시키고 운영하기 위해서는 상당한 비용이 발생한다.

인공 지능 창작물이 인간의 창작물과 동일한 가치를 가지지에 대한 논의가 필요하며, 인공 지능이 창작 활동을 수행하게 되면, 예술가와 디자이너 등 창작자들의 일자리에 영향을 줄 수 있다.

프레히트는 기술의 임무는 유용성이며, 행복 증진이나 의미 창출은 결코 앞자리에 서지 못한다. 그건 인공 지능에도 해당된다[6]. 이는 인공지능 기술의 활용에 있어 유용성과 지속 가능성을 우선시해야 함을 강조한다.

인공지능은 인간의 복잡하고 다양한 감정과 세세한 뉘앙스를 완벽하게 이해하거나 표현하기 어렵다. 또한, 특정 문화나 사회적 맥락을 완전히 이해하거나 반영하는 것 역시 어렵다. 이와 같은 문제점과 한계들은 인공 지능 기술이 예술 분야에서 더욱 발전하고 성숙해지기 위해 극복해야 할 중요한 과제로 여겨진다. 이로써 인공 지능과 인간 예술가가 보다 건설적으로 협력하며 새로운 창작의 영역을 탐색할 수 있을 것이다.

인간과 인공지능의 차이가 중요한 것이 아니라 인공지능을 어떻게 다룰 것인지가 중요하다. 아이디어가 풍부한 작가가 뛰어난 작품을 만들 때 인공지능은 아주 좋은 도구가 될 것이다. 마지막 선택은 예술가의 몫일 것이다 [7]. 인공지능을 예술 창작의 도구로 활용하는 데 있어 작가의 역할과 선택을 시사한다. 인공지능의 한계를 인정

하고, 그것을 보완할 수 있는 인간 예술가의 창의성과 감성이 결합될 때, 비로소 인공지능 기반 자동 창작이 새로운 예술적 지평을 열어갈 수 있을 것이다.

또한 인공지능의 작품을 인간화해서 그로부터 무언가를 느끼는 것도 인간다움의 발현이다. 인간은 세상과 대상에 자신을 투영해서 인간을 읽어낸다. 세상과 대상에 없는 그 무언을 인간이 부여하고 창출 한다[7]. 이는 인공지능 창작물을 대하는 인간의 태도와 인식 자체가 인간다움의 한 측면임을 보여준다. 비록 인공지능이 만들어낸 작품이라 할지라도 감동과 의미를 찾아내는 것은 결국 인간의 몫이다. 이러한 관점에서 보면, 인공지능 기반 자동 창작은 인간의 창의성과 감성을 확장하는 새로운 계기가 될 것이다.

## IV. 결론

인공지능의 급격한 발전으로 창작과 학습 과정에서 인공지능과의 협업에 대한 가능성을 확인할 수 있다[8]. 이는 예술 창작 방식에 혁신을 가져왔으며, 추상적 형태의 예술 작품과 표현 방식을 탐구하는 가능성을 제시한다. 그러나 이러한 기술적 발전은 창의성의 본질, 데이터 보호, 도덕적 책임, 개인에 미치는 영향, 사회적 가치 등 다양한 문제를 초래하며 이에 대한 논의가 필요하다.

결국, 인공지능과 창작자 간의 상호 보완적인 관계를 어떻게 구축하고 발전시킬 것인가를 고민하는 것은 매우 중요하다. 창의적 활동은 기술이 제공하는 새로운 도구와 가능성을 최대한 활용하며, 인간의 독창성, 감성, 문화적 맥락을 존중하는 방식으로 창작 활동을 진행해야 할 것이다. 이를 통해, 우리는 기술과 예술이 서로를 발전시키고 새로운 창작의 영역을 탐구할 수 있는 미래를 구축할 수 있을 것이다.

그러나 Mumford는 자동기계들은 기초적 생산 영역을 끊임없이 정복하려 하겠지만, 교육, 재창조, 실험으로 무장한 수공업과 기계 작업이 반드시 균형을 만들어낼 것이다. 수공업과 기계 작업이 없다면 자동주의는 궁극적으로 사회를 엉망진창으로 만들어버릴 테고, 이 자동주의가 더 발전할 경우 사회의 존립은 위태로워질 것이라고 하였다[1]. 이는 자동화 기술의 무분별한 확장이 사회적 위험을 초래할 수 있음을 시사한다.

따라서 인공지능 기반 자동 창작의 발전 과정에서, 인간 예술가의 역할과 가치를 간과해서는 안 된다. 인간의 창의성과 감수성, 윤리적 판단력 등은 여전히 예술 창작

에 있어 핵심적인 요소이다. 인공지능과 예술가가 조화롭게 협업하고, 기술과 예술이 상호 보완적으로 발전할 수 있을 때, 우리는 자동 창작 예술의 긍정적인 가능성을 실현할 수 있을 것이다.

이를 위해 각계각층의 이해관계자들이 적극적으로 소통하고 협력해야 한다. 인공지능 기반 자동 창작 예술 작품의 사회 문화적, 윤리적, 법적 영향에 대한 심도 있는 연구와 논의를 바탕으로, 기술 발전의 혜택을 극대화하면서도 부작용을 최소화할 수 있는 방안을 모색해야 할 것이다. 이러한 노력을 통해, 우리는 인공지능 시대의 자동 예술 창작이 인간의 창의성과 가치를 존중하는 가운데 새로운 지평을 열어갈 수 있기를 기대해 본다.

## References

- [1] Mumford, L. "Technics and civilization". Book World. 2013, p. 75, 558.
- [2] Newhall, B. "The history of photography". Youlhwadang. 2003, p. 39.
- [3] Arthur, A. I. "The artist in the machine". Culture Books. 2022, pp. 188-193.
- [4] Andrew, M. and Erik, B. "Machine, Platform, Crowd". Chungrim. 2018, pp105-110.
- [5] Foster, D. "Generative deep learning". Hanbit Media. 2023, pp. 412-429.
- [6] Precht, R. D. "Künstliche Intelligenz und der Sinn des Lebens". Open Books. 2022, p. 254.
- [7] Kim, J. I. "AI BigBang". Dongasiabook. 2023 pp. 61-63.
- [8] Kim, Hyeonjin, Kim, Yeongjo, Yun, Donghyeon and Lee, Hanjin, "Empirical Research on the Interaction between Visual Art Creation and Artificial Intelligence Collaboration", Journal of the Convergence on Culture Technology(JCCT), Vol. 10, No. 1, 2024, pp. 517-524, doi: 10.17703/JCCT.2024.10.1.517