

이미지 생성형 AI의 인물사진 구현 능력 비교 연구 -챗GPT와 제미나이를 중심으로-

A Comparative Study on Portrait Photography Implementation of Image Generative AI -Focusing on Chat GPT and Gemini-

최국순¹, 김경수^{2*}

KookSeon Choi¹, KyoungSoo Kim^{2*}

요 약

본 연구는 사진예술 영역에서 AI 도구의 실질적 활용 가능성과 한계를 검증하기 위해 오픈AI의 챗 GPT(달리3)와 구글의 제미나이(나노바나나2)를 대상으로, 연구자의 촬영본 인물 사진 2장을 기반으로 한 이미지 생성을 심층적으로 비교 분석하였다. 실험결과는 다음과 같다. 첫째, 기술적 품질 및 배경 구현에서는 높은 출력 해상도와 카메라 노출 색감에서 제미나이가 우수한 성능을 보였다. 특히 야외에서 카메라 노출 특징을 구현하는 것에서 차이가 두드러졌다. 둘째, 감정표현에서 시각적 표현 강도에 챗GPT가 앞섰다. 제미나이는 절제된 표현이 더 많았다. 셋째, 두 AI 모두 후면과 측면 구도 구현에서 한계를 보였다. 학습 데이터의 전면 편향성에 기인하는 것으로 분석된다. 넷째, 두 AI 모두 인물의 구체적 외모를 정확하게 구현하는 것은 한계가 있었다. 특히 챗GPT는 이상화 현상을 보였고, 제미나이는 상대적으로 원본 충실도가 높았으나, 완전한 구현에는 한계가 있었다. 이미지 생성형 AI의 인물사진 구현은 아직 ‘통계적으로 가능한 이미지’이고 ‘작가의 철학적 시선’이 담긴 이미지는 아니다.

핵심어 : 챗GPT(달리3), 제미나이(나노바나나2), 이미지 생성형 AI, 인물사진 구현

Abstract

This study conducted an in-depth comparative analysis of Open AI's Chat GPT(DALL·E 3) and Google's Gemini(Nano Banana 2) to examine the practical applicability and limitations of AI tools in the field of photographic art. Using two portrait photographs taken by the researcher as source material, the study evaluated each model's image generation performance. The findings are as follows. First, in terms of technical quality and background reproduction, Gemini demonstrated superior performance in output resolution and the rendering of camera exposure and color tones. The difference was particularly pronounced in reproducing the characteristics of camera exposure in outdoor settings. Second, in emotional expression, ChatGPT outperformed in the intensity of visual representation, whereas Gemini tended toward more restrained expression. Third, both AI models exhibited clear limitations in reproducing rear and side compositions—a result attributable to the frontal bias present in their training data. Fourth, both models

1 Dept. of Culture Studies, Graduate School, Chonnam National University, Gwangju, Korea [Graduate Student]
e-mail: hyeonseo4660@gmail.com

2 Dept. of Media Art Technology, Graduate School of Culture, Chonnam National University, Gwangju, Korea [Professor]
e-mail: ks@jnu.ac.kr (Corresponding Author)

Received(May 18, 2026), Review Result(1st: June 12, 2026), Accepted(August 7, 2026), Published(August 31, 2026)



© 2026 The Authors. Published by NCIS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

were limited in their ability to accurately reproduce the specific physical features of the subjects. Chat GPT in particular displayed a tendency toward idealization, while Gemini showed relatively higher fidelity to the original; however, neither achieved complete reproduction. In conclusion, the portrait reproduction generated by image-generative AI remains a 'statistically probable image' rather than an 'image imbued with the artist's philosophical perspective.'

Keyword : Chat GPT(DALL·E 3), Gemini(Nano Banana 2), Image Generative AI, Portrait Photography Implementation

1. 서론

시각 예술의 역사는 기술적 진보와 궤를 같이하며, 과거 초상화에서 사진술로의 이행은 재현의 패러다임을 근본적으로 변화시켰다 [1][2]. 이후 아날로그 필름에서 디지털 카메라로의 전환은 이미지의 생성의 속도와 효율성을 혁신했으며, 포토샵과 같은 편집 툴의 등장은 후보정을 통한 창작의 범주를 무한히 확장했다. 이제 기술은 텍스트 기반의 생성형 인공지능 시대로 접어들며, 이미지 생성의 주체를 인간에서 알고리즘으로 이동시키는 새로운 변곡점을 맞이하고 있다 [3].

특히 챗GPT와 제미니와 같은 멀티모달 AI는 정교한 이미지 변환 및 생성 능력을 통해 사진 예술의 전통적 가치와 제작 방식에 새로운 화두를 던진다 [4-6].

사진가로서 연구자는 이러한 변화 속에서 AI가 이미지의 본질적 일관성과 서사적 맥락을 어떻게 해석하는지 규명할 필요를 느꼈다. 본 연구는 실제 사진 작품을 원본으로 하여, 대표적인 생성형 AI인 챗GPT와 제미니의 이미지 변환 성능을 비교 분석하고자 한다. 인물의 감정 표현과 두 모델이 보여주는 이미지 재구성 방식의 차이를 정밀하게 고찰한다. 또한 원본의 정체성을 보존하는 '편집적' 기능과 새로운 '생성적' 기능의 구조적 특징을 대조하여 분석한다. 이러한 실증적 분석은 AI가 사진가의 창의적 조력자로서 어느 정도의 유효성과 한계를 가지는지 검토하는 데 목적이 있다. 첫째, 챗GPT(달리3)와 제미니(나노바나나2)는 인물 감정 변환 및 각 AI가 생성한 이미지의 얼굴 유사도, 신체 비율, 감정 표현의 자연도를 비교한다. 둘째, 각 AI의 전신 이미지 변환 과정에서 원본 사진 유지 정도와 촬영방향 변화 재현 능력을 분석한다. 셋째, 각 AI 이미지 변환의 기술적 차이가 생성 이미지 품질에 미치는 강점과 한계 영향을 논의한다.

2. 연구 방법

2.1 연구 대상

본 연구는 연구자가 직접 촬영한 인물사진 포트레이트 2점을 원본으로 선택한다. 생성형 AI는 챗GPT와 제미니를 사용한다. 첫 번째 인물사진은 ISO 100, F/1.2, 1/400s, Canon EOS 5D Mark IV 촬영, 두 번째 인물사진은 ISO 100, F/14, 1/250s Canon EOS 5D Mark IV 촬영이다.



[그림 1] 인물사진 1과 인물사진 2

[Fig. 1] Portrait 1 and Portrait 2

2.2 연구 방법과 절차

생성형 AI는 챗GPT의 달리3 [4][5]와 제미나이의 나노바나나2 [7]를 사용한다. 연구의 1단계는 인물사진의 AI 감정표현을 생성하기 위해 희노애락(喜怒哀樂)을 프롬프트로 사용한다. 연구의 2단계에서는 인물사진의 전신 방향표현을 생성하기 위해 전후좌우(前後左右)를 프롬프트로 사용한다. 연구의 3단계에서는 1단계, 2단계의 AI 생성 사진 이미지의 결과물을 비교 분석한다.

[표 1] 실험 방법 및 절차

[Table 1] Experimental Methods and Procedures

인물사진 2점 선택	1단계	2단계	3단계
동일 프롬프트	희노애락(喜怒哀樂)	전후좌우(前後左右)	AI 사진 이미지 비교
AI 인물사진 이미지 생성	인물의 감정 표현 능력	전신의 기술적 구현 능력	챗GPT와 제미나이의 강점 및 한계점 도출

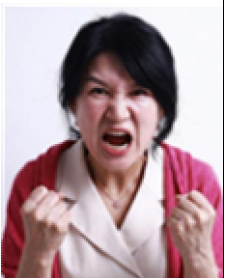
3. 이미지 생성형 AI의 인물사진 구현 능력 비교

3.1 1단계 인물사진의 감정표현 AI 이미지 생성

[그림1](A) 챗GPT와 제미나이의 희노애락 AI 이미지 생성 비교 결과는 [표 2]와 같이 정리하였다.

[표 2] 감정표현의 챗GPT와 제미니의 AI 이미지 생성 결과

[Table 2] Emotional Expression Chat GPT and Gemini AI Image Generation Results

구분	챗GPT	제미니	감정 표현 능력	기술 구현 능력
희			-챗GPT는 기쁨의 이미지가 활기찬 모습, 눈이 커지고, 턱선 가름해짐. 상의 바뀜. 프로필 사진 용도에 맞추어 변환. -제미니는 원본의 훼손 없이 90% 이상 표현. 원본보다 배경과 전반적 톤이 다소 어둡게 표현.	-챗GPT는 확산광으로 변환, 그림자 삭제, 노출을 1/2스톱 올린 하이키 스타일. 화이트 밸런스를 쿨톤 정리. -제미니는 원본의 핑크색의 텍스처의 세부 특징이 충실히 보존, 노출 1스톱 다운.
노			-챗GPT는 과장된 표현 경향이 생성, 미간의 주름을 만들고 찡그린 눈매와 입 모양으로 화가 난 모습 구현. -제미니는 얼굴 톤을 붉게 만들어 강하게 분노 표현, 손 근육 표현의 세부 묘사가 생성.	-챗GPT는 피부보정이 되었으며, 입안의 어두운 부분과 치아 경계를 보정. Contrast를 강조하였으며 시선이 분산되지 않도록 경계정리.-제미니는 생성된 손의 둘레에 배경이 제거되어, 조화와 조형 불균형.
애			-챗GPT는 눈가의 물기와 눈물 표현. 입, 눈 모양을 아래로 처지게 하고 미간의 주름을 만들어 감정 구현. -제미니는 양손의 위치가 다르게 생성되고 충혈된 눈, 정서적 표현 강도가 상대적 증폭.	-챗GPT는 피부보정과 턱선 리터칭. 노출 값이 희, 락 보다 낮아졌으며 대비가 경미하게 조정. -제미니는 노출이 1/3스톱 하향 조정 cool tone 정리. 마젠타 톤 조정으로 감정 강화. 양손 생성.
락			-챗GPT는 생동감이 상승. '희'와의 차이가 불명확. 상의가 원본 대비 이탈이 관찰. '희'와 흡사한, 같은 모양의 손이 생성. -제미니는 표정이 자연스럽게 생성, 원본 유지율이 더 우수.	-챗GPT는 노출이 한 스톱업, 톤 보정과 피부결을 보정, 조리개를 적당히 조여 피사계 심도를 확보. 눈동자의 캐치라이트를 선명하게 처리. -제미니는 핑크색의 채도를 높였고, 화이트 밸런스를 투명하게 보정.

[그림1](B) 챗GPT와 제미니의 희노애락 AI 이미지 생성 비교 결과는 [표 3]과 같이 정리하였다.

[표 3] 감정표현의 챗GPT와 제미나이의 AI 이미지 생성 결과

[Table 3] Emotional Expression Chat GPT and Gemini AI Image Generation Results

구분	챗GPT	제미나이	감정 표현 능력	기술 구현 능력
희			-챗GPT는 기쁜 감정은 잘 표현 되었으나, 원본의 윤곽선 및 구강 구조가 이상화되는 경향. -제미나이는 원본 충실도가 더 높아 주름과 구강 구조의 구현 정확도가 더 높으며 상대적으로 우수.	-챗GPT는 cool tone으로 화이트 밸런스가 상향, 피부 리터칭이 과도하고 원본 일치도 저하. -제미나이는 warm tone으로 원본과 비교하여 톤이 부드럽게 변환. 원본 유지율이 더 높고 자연스럽게 생성처리.
노			-챗GPT는 채도가 낮아지고 어두움으로 분노를 표현, 얼굴 주름이 자연스럽게 생김이 많아 왜곡이 생긴 사각형으로 생성. -제미나이는 원본 유지는 잘 처리되었으나 분노 강도가 낮고 '노'보다 약간의 찡그림으로 생성 처리.	-챗GPT는 명도 조절이 좀 어둡게 표현하여 분노를 표현, 노출이 한 스톱 정도 다운, 구조적 왜곡이 현저. -제미나이는 cool tone, 1/3스톱 정도 상향되어 표정의 디테일 유지. 미세하게 화각이 좁아졌고 손 상향 생성.
애			-챗GPT는 다소 슬픈 감정 생성에 어려움을 보여 찡그림으로 구현, 슬픔이라기보다 약간의 언짢음 정도의 감정이 생성. -제미나이는 눈물이 생성되어 완성도에서 더 높았으나 강한 슬픔에 이르는 생성에는 부족.	-챗GPT는 cool tone을 1/2 스톱 정도 상향 조정, 얼굴형이 원본대비 넓고 라인이 변형. -제미나이는 화이트 밸런스에는 변화가 없으나 눈물방울을 생성, 슬픈 감정 표현. 섬세한 시선, 우측 조명 노출 1/3 스톱 하향 강조.
락			-챗GPT는 '하'와 크게 차이가 없으나 역시 얼굴의 주름이 많이 보정되어 원본의 느낌이 사라지고 이상화 처리. -제미나이는 '하'보다 cool tone이 1/2 스톱 정도 다르게 구현 생성되어 경쾌함이 드러나게 표현.	-챗GPT는 warm tone, glow 효과로 부드러운 빛이 번지는 효과, 변환. 얼굴 라인이 원본에 비해 변형되어 생성. -제미나이는 명암의 대비를 낮추어 질감을 시각적으로 완화, 즐거움을 표현 생성.

3.2 2단계 전신 사진의 AI 이미지 생성

[그림1](A) 전신의 전후좌우 AI 이미지로 생성한 비교 결과는 [표 4]와 같이 정리하였다.

[표 4] 전후좌우의 챗GPT와 제미나이의 이미지 생성 비교 결과

[Table 4] Image generation results of Chat GPT and Gemini in the shooting direction

구분	챗GPT	제미나이	방향 표현 능력	기술 구현 능력
전			-챗GPT는 인물의 키와 체형 비율에서 차이가 있고, 특히 원본 대비 길어지는 경향이 관찰. -제미나이는 원본과 같은 해상도가 유지되었으나 의상과 디테일이 차이가 나고 특히 상의가 명도 다운 처리.	-챗GPT는 가로세로 비율이 2:3, 전체 노출이 1/3 스톱 업 처리. -제미나이는 가로세로 비율이 1:1.23(약4:5), 거의 중형 카메라 비율이 나왔으며 상의 컬러가 원본대비 1/3 스톱 다운되어 생성.
우			-챗GPT는 완전한 우측이 아닌 3/4 각도를 구현. -제미나이도 완전한 우측 모습은 아니며 옷 컬러가 자주 빛으로 바뀌었고, 바닥 컬러가 그레이 톤으로 생성되어 원본의 컬러 유지력과 정합성 미확보.	-챗GPT는 가로세로 비율이 2:3(1:1.5)이 나오고, -제미나이는 가로세로 비율이 1:1.25(4:5)으로 중형 카메라에 근접한 비율 생성. 특히 원본과의 컬러 일치도가 저하, 명도와 콘트라스트의 변화가 있고, 원본대비 이탈이 관찰.
좌			-챗GPT는 얼굴에서 경미한 비대칭이 발견, 정 좌의 모습 구현에 한계 발생. -제미나이는 옆모습 구현에 챗GPT보다 상대적으로 우수하나, 완전한 구현에는 역시 한계. 3/4 각도 생성.	-챗GPT는 가로세로 비율이 2:3이 나오고, -제미나이는 챗GPT와 가로는 동 일하나, 세로가 47픽셀 줄었고, 비율이 5:6(1:1.2) 정도 해당, 챗GPT 사진에 비해 정방형(1:1)에 가까운 형태.
후			-챗GPT는 얼굴이 보이지 않음에도 헤어 스타일과 의상 실루엣의 일관성 유지. -제미나이도 모두 상당부분 유지. 그러나 원본의 컬러, 채도, 대비 등 복장 구현에는 한계 점이 관찰.	-챗GPT는 가로세로 비율이 2:3(1:1.49) 생성. -제미나이는 5:6(1:1.17)로 신체 비율 생성에 한계점이 관찰. 원본대비 이탈이 발생. 약 f/5.6~f/11의 조리개 값으로 보이나 실제 값은 f/14.

[그림1](B) 전신의 전후좌우 AI 이미지로 생성한 비교 결과는 [표 5]와 같이 정리하였다.

[표 5] 촬영 방향 전후좌우의 챗GPT와 제미니 이미지 생성 비교 결과

[Table 5] Image generation results of Chat GPT and Gemini in the shooting direction

구분	챗GPT	제미니	방향 표현 능력	기술 구현 능력
전			-챗GPT는 세로 비율이 길고, 인물 비율면에서 경미한 왜곡이 발생. 우산의 원형이 일부 트리밍 처리. -제미니는 원본에 색감과 구도, 카메라의 노출 설정에 맞게 배경을 아웃 포커싱 처리. 원본 유지율 상대적으로 우수.	-챗GPT는 가로83px 세로159px, 비율이 1:1.91로 표준사진보다 세로로 긴 형태, 스마트폰 풀 스크린 비율 9:16보다 더 좁고 수직적인 느낌. -제미니는 가로122px 세로159px, 비율 1:1.3 카메라의 표준비율 3:4(0.75)에 근접, 안정적 구조.
우			-챗GPT는 우산의 지속적 등장은 확인, 우산의 기울기와 손잡이 구현이 어렵고, 정방향 우측 재현 불가, 3/4 방향 생성. -제미니는 우산이 일관되게 유지, 정우측 방향 구현 불가, 전체적 색상이 Warm tone로 조정.	-챗GPT 비율은 위의 '화'와 동일, 화이트 밸런스에서 챗GPT는 Cool tone으로 조정, 스마트폰 풀 스크린 비율 9:16보다 더 좁고 긴 형태, 수직적인 느낌 강조. -제미니는 warm tone, 원본의 렌즈 85mm, f값 1.2의 효과를 제한적으로 생성.
좌			-챗GPT는 인물이 경미하게 왜곡되고 수직적으로 이상화 생성. -제미니는 카메라 설정 노출값이 일관되게 유지되고 자연스러우나, 정 좌는 챗GPT와 제미니 모두 배경 및 인공물 발생 일관성에 한계 관찰.	-챗GPT는 비율이 1:2, 제미니는 1:1.26(4:5) 정도의 비율. -제미니의 색온도 warm tone, 편안한 색감으로 구현. 85mm 렌즈 장점인 역광의 부드러움을 생성 도출.
후			-챗GPT는 우산의 원형이 심하게 왜곡, 비율도 수직적으로 길게 이상화되어 생성. -제미니는 꽃과 바닥, 담장 등 생성된 인공물, 즉 뒷 배경의 일관성이 떨어지고 각각 정 후면 방향 이행도 저하.	-챗GPT의 비율은 약1:2, 해상도는 가로101, 세로201px. 100% 후면이 정확하게 구현 불가. -제미니는 비율 4:5(1:1.26), 해상도 가로138, 세로174px로 85mm렌즈의 역광 촬영 장점을 제한적 모방.

3.3 챗GPT와 제미니의 강점 및 한계점

본 연구에서 [그림1](A)(B) 감정표현의 희노애락과 촬영방향 전후좌우의 전신, 챗GPT와 제미니의 AI 이미지 생성 종합결과는 [표 6]과 같이 정리하였다.

[표 6] 챗GPT와 제미니의 AI이미지 생성 결과물에 대한 종합 비교 결과

[Table 6] Comprehensive Comparison Results of AI Image Generation Outputs by Chat GPT and Gemini

구분	챗GPT		제미니	
	감정,방향 표현 능력	기술 구현 능력	감정,방향 표현 능력	기술 구현 능력
1단계	감정 표현의 강도가 다소 우세하나, 비율과 대칭에서 왜곡이 다소 심하게 생성. 감정은 감정 표현이 대체적으로 우세하나, 원본 유지력이 떨어지며 얼굴 라인과 피부톤, 구강 및 치아 구조까지 이상화하여 최종 결과물의 완성도가 원본대비 정합성 미흡.	카메라의 노출값과 조리개 값을 한스톱 정도 상향시키며, 피부톤과 색감을 이상화시켜 원본 훼손이 매우 심각. 비율이 비이성적으로 생성되어 원본 유지율이 현저하게 하락.	원본과 비교하여 자연스럽고 큰 훼손은 없는 편이나, 노출값이 변하여 어두워지거나 밝아지는 등 일관성이 부족하고, 소품이나 의상의 유지가 되지 않으며, 추정으로 지시에 임하고 생성된 경향이 다수 발견.	출력 해상도가 유의미 1856×2344px. 챗GPT는 최대 1131×1390px로 강점은 높은 출력 해상도 유지, 노출값이 한스톱 정도 다운되어 다시 밝기 조정이 필요한 상황 발생. 원본의 인물이 다소 유지되고 있으나, 의류나 소품에서 다른 모습이 구현되는 문제점.
2단계	좌우, 후면까지 정각도가 재현불가, 좌우 3/4방향 모습으로 구조적 왜곡이 현저하게 발생, 원본구현 한계가 관찰.	비정상적으로 비율이 차이가 나며(1:2), 이에 우산과 우산 손잡이에서 왜곡이 심하게 발생, 지시성에 정확하게 따르지 못하고, 노출의 일관성 미흡.	전반적으로 안정감 있으며 warm tone으로 편안한 느낌의 구현이 되었으나, 전후좌우의 정확한 방향 자세는 재현 불가. 또한 비율이 다소 변하여 중형 카메라 비율이 발생.	방향 지시성에는 완전하게 따르지 못하였으나, 카메라의 노출과 조리개 값은 원본 유지력이 챗GPT에 비하여 우수. 비율과 픽셀 유지가 높았고 촬영각도의 빛의 부드러움 일관되게 유지.

종합결과의 특징은 ①후면, 즉 뒷 모습에서 두 AI 모두 낮은 구현 충실도를 보였다. 이는 일반적인 이미지 데이터가 정면 및 3/4 각도 인물 사진에 편향되어, 뒷모습 이미지 표본이 상대적으로 부족한 이유로 추정된다 [8][9]. ②옆면, 완전한 측면(90°)구현의 어려움을 보였다. 두 AI 모두 완전한 측면보다 3/4 각도 이미지를 생성했다. ③향후 연구 방향으로는 더 다양한 모델, 촬영 조건 포함한 확장표본 실험이 필요하며, 정량적 유사도 지표 도입, 다수 평가자에 의한 블라인드 평가, 인물의 아이덴티티 보존을 위한 파인튜닝 기법과의 비교 연구도 필요하다. 두 플랫폼 모두 빈번한 업데이트가 이루어지는바, 본 결과는 해당 시점의 단면적 평가라는 시간적 한계가 존재한다.

두 AI의 결과 차이점은 모델의 구조적 특성에 기인하는 것으로 해석된다. 챗GPT는 텍스트 기반

의 이미지 생성에 최적화되어 감정, 동작 등 의미적 지시해석에는 강점을 보이거나, 입력 이미지 정보를 활용하는 과정에서 원본의 픽셀정보가 재해석되어 이상화 현상이 발생했다 [4][10]. 반면 제미나이는 이미지 작업에 더 특화되어 원본의 픽셀정보, 색감, 구도를 보존하는데 강점을 보이지만 감정과 같은 추상적 지시에는 보수적으로 작동했다. 이러한 차이가 이 연구에서 관찰된 강점과 한계의 분기를 만든 것으로 추정된다.

4. 결론

본 연구는 챗GPT와 제미나이의 이미지 생성 능력, 감정표현 희노애락과 전신방향 전후좌우의 생성, 두 실험군에서 비교 분석하였고, 결론은 다음과 같이 요약되었다. 첫째, 기술적 품질 및 배경 구현에서는 높은 출력 해상도와 카메라 노출 색감에서 제미나이가 우수한 성능을 보였다. 특히 야외에서 카메라 노출 특징을 구현하는 것에서 차이가 두드러졌다. 둘째, 감정표현에서 시각적 표현 강도에 챗GPT가 앞섰다. 제미나이는 절제된 표현이 더 많았다. 셋째, 두 AI 모두 후면과 측면 구도 구현에서 한계를 보였다. 학습 데이터의 전면 편향성에 기인하는 것으로 분석된다. 넷째, 두 AI 모두 인물의 구체적 외모를 정확하게 구현하는 것은 한계가 있었다. 특히 챗GPT는 이상화 현상을, 제미나이는 상대적으로 원본 충실도가 높았으나, 완전한 구현에는 한계가 있었다. 그동안 시대마다 기술과 사회적, 문화적 상호작용이 ‘사진이란 무엇인가’를 재정의해 왔다. 인간의 창의성은 단순히 대체하는 것이 아니라, 의미있는 방향으로 확장되어 왔다. AI가 생성한 이미지들은 명백한 한계가 드러나며 [1][2][10], 이는 ‘통계적으로 가능한 이미지’이지 ‘작가의 철학적 시선’이 담긴 이미지는 아니다. 카메라의 부재와 광학의 법칙이 없고, 이미지 이전이 존재하지 않는다. 그럼에도 인간의 직관과 기계학습을 결합하는 새로운 하이브리드 예술 형식을 창작의 증강도구로 삼아 계속 나아갈 것이다 [11]. 이미 수백 번의 프롬프트 작업을 통해 다양한 창작 형태가 등장하고 있는 만큼, 앞으로 작가들이 어떠한 방향과 형태로 AI와 동행하는지 주의 깊게 지켜볼 필요가 있을 것이다.

종합적으로 활용 목적에 따른 도구 선택이 중요하며, 창작조력의 일정부분 도움을 얻을 수는 있겠으나 한계점은 분명했다. 또한, 향후 두 플랫폼의 지속적인 업데이트와 기술의 발전에 따라 이러한 격차는 변화할 것이므로 앞으로도 비교연구가 필요할 것이다.

References

- [1] R. Barthes, *Camera Lucida: Reflections on Photography*, R. Howard, Trans., Hill and Wang, 1981.
- [2] S. Sontag, *On Photography*, Farrar, Straus and Giroux, 1977.
- [3] A. Oksanen et al., "Artificial intelligence in fine arts: A systematic review of empirical research," *Computers in Human Behavior: Artificial Humans*, vol. 1, no. 2, Art. no. 100004, Aug.-Dec. 2023, doi: 10.1016/j.chbah.2023.100004.
- [4] J. Betker et al., "Improving image generation with better captions," OpenAI, San Francisco, CA, USA, Oct. 2023. [Online]. Available: <https://cdn.openai.com/papers/dall-e-3.pdf>
- [5] OpenAI, "DALL·E 3 system card," *openai.com*, <https://openai.com/index/dall-e-3-system-card> (accessed Apr. 30, 2026).
- [6] Google DeepMind, "Introducing Nano Banana Pro: Gemini 3 Pro Image," *blog.google*, <https://blog.google/innovation-and-ai/products/nano-banana-pro> (accessed Apr. 30, 2026).
- [7] Google DeepMind, "Nano Banana 2: Closing the gap between speed and beauty (Gemini 3.1 Flash Image)," *blog.google*, <https://blog.google/innovation-and-ai/technology/ai/nano-banana-2> (accessed Apr. 30, 2026).
- [8] A. Borji, "Generated faces in the wild: Quantitative comparison of Stable Diffusion, Midjourney and DALL-E 2," Oct. 2022, arXiv:2210.00586, doi: 10.48550/arXiv.2210.00586.
- [9] A. A. Novikov, M. Vranka, F. David, and A. Voronin, "Can text-to-image generative models accurately depict age? A comparative study on synthetic portrait generation and age estimation," Feb. 2025, arXiv:2502.03420, doi: 10.48550/arXiv.2502.03420.
- [10] C. B. Horton Jr., M. W. White, and S. S. Iyengar, "Bias against AI art can enhance perceptions of human creativity," *Scientific Reports*, vol. 13, no. 1, Art. no. 19001, Nov. 2023, doi: 10.1038/s41598-023-45202-3.
- [11] J. Matthews, D. Fastnedge, and A. Nairn, "The future of advertising campaigns: The role of AI-generated images in advertising creative," *Journal of Pervasive Media*, vol. 8, no. 1, pp. 29-49, Aug. 2023, doi: 10.1386/jpm_00003_1.