

생성형 AI기반 웹툰 제작의 일관성 제어 모델 연구

A Consistency Control Model

for Webtoon Production Based on Generative AI

박희운¹

Hee-Woon Park¹

요약

본 연구는 생성형 인공지능(Generative AI)의 도입이 웹툰 제작 전 과정에 미치는 영향과 시각적 일관성 확보 메커니즘을 분석하는 것을 목적으로 한다. 웹툰 제작은 캐릭터 디자인, 배경 구현, 컷 연출 등 복합적인 시각 요소가 다 같은 서사적 연속성을 유지해야 한다는 점에서 높은 노동 강도와 반복 작업을 요구한다. 본 연구는 Stable Diffusion, ControlNet, Lora 등 생성형 이미지 모델을 활용하여 웹툰 제작 공정을 단계별로 구성하고, 캐릭터·배경·스토리 전개에의 시각적 일관성을 유지하기 위한 기술적 및 디렉팅적 개입 요인을 인간과 생성형 AI의 협업 제작 모델의 가능성을 추론 하였다. 그 결과, 생성형 AI 분석 결과, 생성형 AI는 반복적 시각 작업의 효율과 형태·환경적 일관성 유지에는 효과적이거나, 감정 표현과 서사적 흐름 조율은 인간의 해석과 편집 판단이 핵심적으로 작용함을 확인하였다. 본 연구는 웹툰 산업에서 생성형 AI를 보조 제작 체계로 도입 시 고려해야 할 일관성 관리 전략을 제안함으로써, 향후 인간과 생성형 AI 협업 기반 창작 생태계 구축에 기초를 제공한다.

핵심어 : 웹툰제작, 생성형AI, 일관성, 디렉팅, 제작과정

Abstract

This study aims to analyze the impact of Generative AI on the overall webtoon production process and to identify mechanisms for maintaining visual consistency. Webtoon production requires high labor intensity and recurring visual work, as multiple visual components-such as character design, background rendering, and panel composition-must consistently support the continuity of the narrative. In this study, Stable Diffusion, ControlNet, and LoRA were applied to structure the production workflow in sequential stages and to examine both technical and directing-based interventions necessary to sustain visual coherence in character appearance, background tone, and narrative progression. The findings indicate that generative AI is effective in reducing repetitive visual tasks and maintaining morphological and environmental consistency; however, emotional expression and narrative rhythm require human interpretation and editing decisions. This study suggests strategic guidelines for integrating generative AI as a supportive production system within the webtoon industry and provides a foundation for developing future human-AI collaborative creative ecosystems.

Keyword : Webtoon Production, Generative AI, Consistency, Directing, Production Pipeline

¹ Department Design, Kyungnam University, Kyungnam Korea [Professor]
e-mail: design@hkyungnam.ac.kr

* 이 논문은 2025년도 경남대학교 학술연구비의 지원을 받아 연구되었음.

Received(November 13, 2025), Review Result(1st: December 8, 2025, 2nd: January 22, 2026), Accepted(March 13, 2026), Published(March 31, 2026)



© 2025 The Authors. Published by NCISS.
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/>.

1. 서론

웹툰 산업은 디지털 플랫폼의 확대와 모바일 중심의 콘텐츠 소비 패턴 변화에 따라 빠르게 성장해왔다 [1]. 모바일 화면의 스크롤 기반의 수직형 서사 구조는 독자가 장면을 시간의 흐름 속에서 따라가도록 하여, 기존 만화와는 다른 감각적 몰입을 제공한다 [2]. 이로 인해 한국 웹툰은 글로벌 시장에서 경쟁력을 확보하였지만, 제작 과정은 여전히 기획·작화·배경·색채·컷 연출 등 세부 공정이 촘촘하게 분업된 노동집약적 구조를 유지하고 있다 [3]. 이러한 구조는 장기 연재 속도 유지의 부담, 작가 개인 역량 의존성, 퀄리티 관리의 불균형과 같은 문제를 지속적으로 발생시킨다. 최근 생성형 인공지능 기술의 발전은 웹툰 제작 과정의 일부 자동화 가능성을 제시하고 있다. Stable Diffusion과 같은 이미지 생성 모델은 텍스트 기반 스타일 표현이 가능하며, ControlNet·LoRA·Sref 등의 보조 모델은 구도·형태·색감·재질의 반복성을 일정 수준 통제할 수 있도록 한다 [4]. 즉, 생성형 AI는 배경 제작, 스타일 유지, 장면 재구성 등 반복적 시각 작업의 부담을 완화하는 역할을 수행할 잠재력을 갖는다. 그러나 웹툰은 단일 이미지가 아니라 연속되는 장면들의 서사 흐름 속에서 감정과 의미가 전달되는 매체이다 [5]. 따라서 중요한 것은 이미지의 고품질 자체가 아니라, 컷과 컷, 회차와 회차 사이에서 감정·표정·비례·톤이 일관되게 유지되는가이다. 즉, 생성형 AI 도입에서 핵심은 무엇을 자동화할 것인가·보다 작품의 시각적 정체성과 서사적 연속성을 어떻게 유지할 것인가·의 문제이다. 이에 본 연구는 웹툰 제작 공정을 단계별 파이프라인 구조로 명확히 정의하고, 각 단계에서 발생하는 시각적 일관성 유지 체계를 구조적으로 분석한다. 특히 생성형 AI가 담당할 수 있는 생성 영역과, 인간 디렉터가 개입하여 서사 의미·감정 리듬·연출 강조점을 통제해야 하는 영역을 명확히 구분한다. 즉, 본 연구의 핵심 기여는 AI로 그리는 것이 아니라, AI가 생성하는 결과를 어떻게 지휘하고 해석하며 조정할 것인가를 규정하는 데 있다. 이는 향후 웹툰 제작 환경에서 창작 효율성과 작품 정체성 유지의 균형을 확보하기 위한 실질적 공정 재설계 모델을 제공한다는 점에서 의미를 갖는다.

2. 관련 연구 고찰

2.1 생성형AI와 웹툰제작

생성형 AI 기술은 이미지 생성의 자율성과 표현 가능 범위를 확장시키며 시각 콘텐츠 제작 전반에 영향을 미치고 있다. Stable Diffusion과 같은 텍스트-이미지 생성 모델은 LoRA, ControlNet 등 제어 모듈과 결합되면서 스타일 유지, 구도 조절, 캐릭터 반복 표현이 가능해졌다 [6]. 그러나 이러한 발전은 단일 이미지의 품질 향상에 주로 집중되어 있으며, 장면 간 서사 흐름이 중요한 웹툰에

서 이미지가 어떤 맥락 속에서 연결되고 해석되어야 하는지에 대한 논의는 충분히 이루어지지 않았다 [7]. 웹툰 제작은 기획, 콘티, 캐릭터 설정, 배경 구현, 컷 연출 등 단계적으로 전개되며, 독자는 캐릭터 외형·장면 분위기·감정 흐름이 연속적으로 유지될 때 하나의 세계를 경험한다. 본 연구는 이를 시각적 일관성이라 정의하며, 이를 형태적·서사적·환경적·인지적 측면으로 구분한다. 그러나 기존 연구는 주로 배경 자동화, 스타일 변환 등 개별 작업 효율 향상에 치중하여, 장면과 장면이 연결되는 서사 미디어로서의 웹툰 특성을 충분히 반영하지 못했다. 또한 생성형 AI에 대한 협업 모델 논의는 인간과 AI가 생성(G)-선택(S)-조정(D)의 순환 구조로 역할을 나누어야 한다는 방향으로 진행되고 있으나, 웹툰 제작 공정에서 각 단계별 역할 분담이 명확히 제시된 연구는 드물다. 이에 본 연구는 생성형 AI를 웹툰 제작 파이프라인에 적용할 때 발생하는 시각적 일관성 문제를 공정 단위에서 체계적으로 분석하고, AI가 담당할 수 있는 생성 영역과 인간 디렉터가 개입해야 하는 조정 영역을 명확히 구분하고자 하는 것이 연구의 목표이다.

2.2 웹툰 제작 구조의 특징

웹툰 제작은 단순한 이미지 생성이 아니라 서사 의미와 시각 표현이 결합되는 통합적 창작 과정이다 [6]. 제작 공정은 ① 기획 및 콘티, ② 캐릭터 스타일 설정, ③ 배경 구성, ④ 컷 연출, ⑤ 후반 편집의 단계로 이루어지며, 작가-어시스턴트-채색-배경 담당자 등 분업적 협업 구조를 기반으로 운영된다 [7]. 주간 연재 체계에서 각 회차는 독립적 이미지가 아니라 연속되는 서사 단위로 작동하므로, 작가는 빠른 작업 속도 속에서도 캐릭터 외형, 감정 흐름, 배경 톤을 안정적으로 유지해야 한다. 또한 웹툰은 세로 스크롤 기반 표현 형식을 가지며, 컷 간 간격, 클로즈업 타이밍, 장면 전환 속도는 독자의 몰입과 감정 해석 리듬을 조절하는 연출 요소로 기능한다. 이 때문에 [표 1]에서 정리한 내용을 보면 시각적 일관성은 ① 형태적, ② 서사적, ③ 환경적, ④ 인지적 측면에서 유지될 필요가 있다. 따라서 생성형 AI를 웹툰 제작에 적용할 때 핵심은 단일 이미지의 완성도가 아니라, 컷·장면·회차 전반에 걸친 연속성과 정체성 유지이다. 즉, 자동 생성만으로는 충분하지 않으며, 이를 조율·관리하는 일관성 제어 모델이 요구된다.

[표 1] 웹툰의 일관성 유지의 내용과 의미

[Table 1] Maintaining Consistent Content and Meaning in Webtoons

차원	내용	웹툰에서의 의미
형태적 일관성	캐릭터 외형·색·질감의 반복성	동일 인물로 인지되는지 여부
서사적 일관성	감정 흐름 및 관계 전개 연속성	이야기 전개가 자연스럽게 연결되는지
환경적 일관성	배경 톤·공간 구조의 통일	세계관의 안정감 유지
인지적 일관성	독자의 시선 흐름과 정보 해석의 안정성	스크롤 과정에서 몰입이 유지되는지

2.3 생성형 인공지능의 이미지 생성 구조

생성형 AI의 이미지 생성은 대규모 데이터 학습과 확률적 복원 과정에 기반한다 [8]. Stable Diffusion 모델은 텍스트 입력을 의미 벡터로 변환하고, 초기 노이즈를 점진적으로 정제하여 이미지를 생성하는 확산(Diffusion) 구조를 따른다 [6]. 이때 프롬프트는 단순 설명이 아니라 스타일·구도·감정·광원 조건을 포함한 의미 지시 정보로 작동한다. [표 2]에서 정리한 내용을 설명하자면 CLIP 모델은 텍스트와 이미지 간의 공유 의미 공간을 형성하여, 어떤 문장이 어떤 시각 표현과 대응하는지 계산 가능하게 만든다. 결국 생성 결과는 학습 데이터의 통계적 패턴에 기반한 잠재 표현(latent representation)의 재구성이라 할 수 있다 [6]. 웹툰 제작에서는 동일 인물과 장면을 반복적으로 표현해야 하므로, 추가적인 제어 기술이 필요하다. LoRA는 캐릭터의 스타일 속성을 모델 내부에 내재화하여 반복 재현을 가능하게 하고, Sref는 참조 이미지를 기준으로 하여 스타일 일탈을 방지한다 [9]. ControlNet은 포즈·구도·시점 구조를 고정함으로써 장면 구성과 연출 흐름을 유지하도록 한다. 즉, LoRA와 Sref는 어떤 모습으로, ControlNet은 어떻게 배치할 것인가를 제어한다. 그러나 이러한 기술들은 주로 형태와 스타일의 표면적 일관성 유지에 초점을 두며, 감정 표현의 강약, 장면 간 서사 리듬, 관계 변화의 의미 강조와 같은 고차원적 서사적 판단은 AI가 단독으로 수행하기 어렵다. 따라서 생성형 AI는 반복·밀도 높은 시각 작업을 보조하고 톤과 스타일을 유지하는 역할에 적합하며, 감정과 서사 해석은 인간 디렉터의 개입을 필요로 한다. 이에 웹툰 제작에서는 AI 생성 + 인간 해석·조정이 결합된 협업형 일관성 제어 모델이 요구된다는 특징적인 요인을 찾을 수 있었다.

[표 2] AI기술 요소의 웹툰제작 기능과 역할

[Table 2] Webtoon Production Functions and Roles of AI Technology Elements

기술 요소	기능	웹툰 제작에서의 역할
Sref (Style Reference)	특정 스타일 기준점 제공	캐릭터 및 장면 전체 질감 유지
LoRA (Low-Rank Adaptation)	개별 스타일 압축 학습	캐릭터 형태·색채 반복성 확보
ControlNet	구도·포즈·윤곽 제어	서사 맥락에 따른 행동 흐름 유지
LUT 톤 매핑	색채 및 조명 통합	세계관 분위기 일관 유지

3. 연구방법 및 분석

본 연구는 생성형 인공지능을 웹툰 제작 공정에 적용하는 과정에서 시각적 일관성을 유지하기 위한 메커니즘을 규명하기 위해 정량적 분석과 정성적 평가를 병행하였다. 분석 단위는 캐릭터 표현, 배경 스타일, 컷 흐름의 서사 연결이며, AI 자동 생성 결과와 인간 디렉팅 개입이 반영된 결과

를 비교하였다. 정량적 분석에는 SSIM과 CLIP Similarity를 사용하여 형태 및 의미 연속성을 측정하였고, 정성적 평가는 웹툰 제작 경험자와 플랫폼 독자 집단을 대상으로 인지적 일관성에 대한 평가를 실시하였다. 본 연구는 작품 완성도 평가가 아니라, 웹툰 제작 파이프라인에서 AI의 개입 위치와 디렉팅 통제 방식이 일관성에 미치는 영향을 밝히는 것을 목표로 한다.

3.1 시각적 일관성 평가 기준 및 지표 설정

시각적 일관성은 웹툰에서 독자가 동일한 세계와 인물을 지속적으로 인식하도록 하는 핵심 요소이다 [10]. 본 연구는 이를 형태적·서사적·환경적·인지적 측면으로 구분하고, 생성형 AI 도입 시 발생할 수 있는 외형 변화, 감정 흐름 단절, 배경 톤 불일치를 평가하기 위한 지표를 설정하였다. 형태적 일관성은 캐릭터 외형과 색감 유지 정도를 의미하며 SSIM·CLIP Similarity와 전문가 평가로 확인하였다. 서사적 일관성은 컷과 장면 전개 속에서 감정·관계 흐름의 자연성을 정성적으로 평가하였다. 환경적 일관성은 배경·색온도·조명 톤의 지속성을 Mean Color Deviation과 톤 통합 도구로 점검하였고, 인지적 일관성은 세로 스크롤에서 시선 이동 흐름이 부드럽게 유지되는가를 기준으로 분석하였다. 이에 본 연구는 생성형 AI 자동 생성 → 인간 디렉팅 → 최종 편집으로 구성된 일관성 제어 모델을 제안한다. [표 3]과 같이 SSIM은 캐릭터 외형(눈, 얼굴 비례, 헤어 라인 등)의 구조적 안정성을, CLIP Similarity는 동일 서사 지시 하의 의미적 일관성을 평가하는 지표로 활용하였으며 SSIM 0.65 이상, CLIP-Sim 0.28 이상을 일관성 유지 기준으로 설정하였다. [표 4]의 내용으로 정성 평가는 전문가 패널과 독자 인지 평가로 구성되었으며, 전문가 그룹은 감정·사건 흐름의 연속성, 캐릭터 정체성 유지, 장면 전환 자연성, 의도 전달력을 중심으로 검토하였다.

[표 3] 일관성 지표 정량 평가 기준 표

[Table 3] Consistency Indicator Quantitative Evaluation Criteria Table

지표	산출 방식	측정 대상	평가 목적
SSIM (Structural Similarity Index)	두 이미지의 구조·명암 대비·패턴 유사도를 비교	동일 캐릭터의 얼굴, 신체 비례, 특징 요소	형태적 일관성 유지 여부 확인
CLIP Similarity Score	텍스트·이미지 간 의미 정합성 비교	장면 설명(프롬프트)과 생성 이미지 간 의미 대응	서사적·상징적 이미지 일관성 측정

[표 4] 일관성 지표 정성 평가 기준 표

[Table 4] Consistency Indicator Qualitative Evaluation Criteria Table

평가 항목	판단 요소	평가 방식	평가 척도
캐릭터 재현 일관성	표정·비례·의상·선 스타일의 반복 안정성	전문가 검토	Likert 5점
배경/환경 스타일 일관성	건축·풍경·조형미·색채 톤 연결	전문가 검토	Likert 5점

컷 구성/연출의 서사 연속성	장면 간 구도 변화, 감정 흐름의 연결감	독자 평가 + 전문가 평가	Likert 5점
전체 몰입감 및 위화감 유무	독자의 심리적 집중감 유지 여부	독자 평가	Likert 5점

3.2 평가 분석 절차와 해석 방법

본 연구는 생성형 인공지능 기반 웹툰 제작에서 시각적 일관성 제어 메커니즘의 효과를 확인하기 위해 AI 단독 생성 조건과 AI + 인간 디렉팅 개입 조건을 비교하였다. 동일한 캐릭터와 장면 서사를 기준으로 두 조건에서 각각 장면 시퀀스를 생성하고, 이를 컷 단위 배열로 제시하여 실제 웹툰 감상 맥락에서 평가하였다. 평가 과정은 (1) 데이터 생성 → (2) 정량적 유사도 분석 → (3) 정성적 인지 평가 → (4) 종합 해석의 단계로 진행되었다. 먼저 정량 분석에서는 형태 및 색채 유사도 지표를 활용하여 두 조건 간 시각적 안정성과 구조적 일관성을 비교하였다. 이어 정성 분석에서는 전문가 패널과 독자 인지 평가를 통해 장면 전환의 자연성, 감정·서사 흐름의 연속성, 캐릭터 정체성 유지 여부를 검토하였다. 마지막으로 정량 결과와 정성 평가를 통합하여, 생성형 AI가 유지 가능한 일관성의 범위와 인간 디렉션이 개입해야 하는 핵심 조정 지점을 도출하였다. 이를 통해 본 연구는 생성형 AI가 웹툰 제작 과정에서 수행할 수 있는 역할과 한계를 구체적으로 규명하였다.

3.3 AI 생성을 통한 유사도 분석

동일한 스토리 시놉시스와 캐릭터·장면 설정을 바탕으로 Stable Diffusion을 이용해 두 가지 조건에서 이미지를 생성하였다.

[표 5] 웹툰 데이터 생성과 AI 적용

[Table 5] Webtoon Data Creation and AI Application

조건	설명	AI 적용 수준
조건 A (AI 단독 생성)	텍스트 프롬프트만으로 이미지 생성	모델 자동화 중심
조건 B (AI + 디렉팅 개입)	Sref/Lora 기반 캐릭터 스타일 고정 + ControlNet 구도 통제 + 인간 편집자 수정 반영	인간 조정 중심의 협업 생성

[표 5]와 같이 생성형 AI 단독 생성 조건에서는 텍스트 프롬프트만을 입력하여 모델이 장면을 자동 생성하도록 하였으며, 스타일·구도·표정 변화는 모델의 확률적 출력에 맡겼다. 반면 AI + 인간 디렉션 개입 조건에서는 동일한 프롬프트에 LoRA(캐릭터 스타일 유지), Sref(색감·질감 기준 고정), ControlNet(포즈·구도 제어)을 추가하여 시각적 일관성을 강화하였다. 각 조건에서 생성된 결과

는 장면 단위(6~8컷 시퀀스) 로 구성하여 실제 웹툰 감상 맥락에서 평가하였다. 이를 통해 본 연구는 단일 이미지의 품질이 아니라, 컷과 장면이 연속될 때 시각적 일관성이 유지되는가를 검증하였다. 이후 SSIM 및 CLIP 유사도 분석을 수행하고, 정성적 감상 평가 단계로 이어졌다.

3.4 정량적 유사도 분석

정량적 유사도 분석은 생성형 AI가 웹툰 제작에서 형태적·환경적 시각 일관성을 유지할 수 있는지를 확인하기 위한 단계이다. [표 6]에서와 같이 동일 장면을 대상으로 두 조건을 비교한 결과, SSIM 분석에서는 생성형 AI 단독 생성 조건에서 캐릭터 형태 변동이 크게 나타난 반면, 디렉션 개입 조건에서는 외형이 안정적으로 유지되었다. CLIP 유사도 분석에서도 디렉션 개입 조건이 서사 맥락과 감정 표현 측면에서 더 높은 의미 정합도를 보였다. 또한 Mean Color Deviation 결과, 디렉션 개입 조건에서 배경 색조와 조명 분위기의 일관성이 유지되었다. 종합하면, 생성형 AI는 형태·톤 수준의 기본 일관성 유지에는 효과적이거나, 감정과 서사 흐름의 연속성 확보에는 인간 디렉션 개입이 필수적함이 확인되었다.

[표 6] 정량평가 내용의 평가방식 분류

[Table 6] Classification of Evaluation Methods for Quantitative Evaluation Content

평가 차원	평가 목적	사용 지표/도구	분석 내용	해석 기준
형태적 일관성 (Morphological Consistency)	동일 캐릭터로 인지 가능한가 여부 평가	SSIM (Structural Similarity Index)	얼굴·신체 비례, 윤곽, 선형 구조의 유사성 측정	값이 1에 가까울수록 형태 안정성 및 구조적 일관성 높음
의미적 일관성 (Semantic Consistency)	이미지가 장면 설명(프롬프트) 의미에 부합하는가	CLIP Image-Text Similarity	이미지와 텍스트 의미 벡터 간 유사성 비교	값이 높을수록 서사 내용과 시각 표현의 정합성이 우수함
스타일/질감 반복성 (Stylistic Continuity)	작품의 선화 스타일·채색 질감 일관성 확인	CLIP Image-Image Similarity	컷 간 스타일 및 질감 재현 정도 비교	값이 높을수록 컷 간 표현 양식의 통일성 유지됨

3.5 정성적 유사도 분석

정량 지표만으로는 감정 표현이나 장면 몰입 흐름을 충분히 설명하기 어렵기 때문에, 본 연구는 정성적 평가를 병행하였다. 평가는 전문가 패널 5인과 독자 인지 평가 20인으로 구성되었다. 전문가 평가는 웹툰 제작 및 연출 경험을 가진 평가자를 대상으로 하였으며, [표 7]에서와 같이 생성된 장면 시퀀스의 ① 캐릭터 일관성 유지 ② 배경과 스타일 통일성 ③ 장면 서사의 연속성 ④ 장면 몰입감을 검토하였다. 독자 인지 평가는 실제 감상 환경과 동일한 세로 스크롤 방식으로 진행되었으며, 장면 해석 용이성, 감정 몰입 유지, 시선 이동 흐름을 중심으로 평가하였다. 참여자는 몰입이 끊기거나 해석이 어려운 지점을 자기보고 방식으로 기록하였다. 정성 평가는 단순 선호 판단이 아

나라 이미지가 서사적 세계 속에서 하나의 연속된 경험으로 받아들여지는지를 확인하는 절차이다. 평가는 Likert 5점 척도로 수집되었으며, 두 조건 간 평균 차이를 비교하여 인간 디렉션 개입의 일관성 유지 효과를 검증하였다.

[표 7] 정성평가내용의 평가방식 분류

[Table 7] Classification of Evaluation Methods for Qualitative Evaluation Content

평가 구분	평가 대상	평가 주체	세부 평가 항목	평가 방법	지표/수집 방식
전문가 패널 평가 (Expert Review)	장면 시퀀스 전체(6~8컷)	웹툰 작가, 편집 디렉터, 시각 서사 연구자 (5인)	① 감정사건 흐름의 연속성 ② 캐릭터 정체성 유지 ③ 장면 전환 자연성 ④ 연출 의도 명료성	전문가 개별 리뷰 + 토의형 리뷰	Likert 5점 척도 정성 코멘트
독자 인지 평가 (User Perception Test)	동일 장면 시퀀스를 스크롤 방식으로 제시	대학생·웹툰 이용 독자 20명	⑤ 감정 몰입 유지도 ⑥ 장면 해석 명료성 ⑦ 시선 이동 흐름 자연성 ⑧ 컷 간 감정 연결성	실제 감상 환경 재현(모바일 스크롤)	Likert 5점 척도 + 자기보고식 기록
정성 평가 종합 해석	전문가 + 독자 결과 통합	-	①-⑧ 항목 통합 분석	정량 대비 정성 비교	점수 비교 + 내러티브 해석

3.6 종합 분석

본 연구에서 정의한 시각적 일관성은 단일 지표로 환원될 수 있는 개념이 아니라, 형태적 유사성(Structural Similarity), 스타일 질감의 연속성(Stylistic Coherence), 서사적 몰입감(Narrative Immersion)이 복합적으로 유지되는 지각 체계로 이해된다. 먼저 정량 분석 결과, AI 단독 생성 조건의 평균 SSIM 값은 0.58, CLIP-Sim (Image-Text)은 0.24로 나타났으며, 이는 동일 캐릭터 및 장면에서 형태와 색채 톤이 불안정하게 재현됨을 의미한다. 반면 AI + 인간 디렉션 개입 조건에서는 SSIM 평균 0.72, CLIP-Sim 평균 0.31로 유의미한 상승을 보였다(Δ SSIM = +0.14, Δ CLIP = +0.07). 이는 인간 디렉터의 개입이 형태 구조 안정성과 의미 정합도를 모두 개선시켰음을 시사한다. 또한 Mean Color Deviation(MCD) 분석에서도 AI 단독 생성의 평균 편차가 ΔE 14.8, 디렉션 개입 조건은 ΔE 7.2로 나타나, 후자의 색채·조명 톤이 약 51% 더 안정적으로 유지되었다. 그러나 이러한 수치적 안정성만으로는 웹툰 독자가 경험하는 감정 흐름과 서사 연속성을 충분히 설명할 수 없었다. 정성적 평가에서는 전문가 패널(5인)이 평가한 서사적 자연성 평균 4.3/5점, 감정 몰입 유지도 4.1/5점, 캐릭터 정체성 일관성 4.4/5점이 기록되었으며, 이는 AI 단독 생성 조건(평균 3.2~3.4점대)보다 전반적으로 높게 나타났다. 독자 인지 평가(20명) 결과에서도 컷 간 감정 단절 지점이 32% 감소하였고, 장면 해석의 명료도는 27% 향상되었다. 이러한 결과는 [표 8]에 제시된 바와 같이, 정량적 안정성이 시각적 일관성의 “기초 조건”을 제공하지만, 감정선·연출 리듬·관계 맥락 등의 고차원적 연결성은 정성 평가를 통해서만 확인될 수 있음을 보여준다. 즉, 시각적 일관성은 단순히 캐릭터

형태가 동일하게 반복되는 수준이 아니라, 형태의 안정성(quantitative stability) 과 서사의 연속성(qualitative coherence) 이 통합적으로 확보될 때 비로소 ‘지속적 감상 경험(consistency in experience)’ 이 유지된다. 결과적으로, 본 연구는 생성형AI 기반 웹툰 제작에서 정량적 일관성 확보는 AI 시스템이 담당할 수 있는 기술적 영역이며, 정성적 일관성은 인간 디렉팅이 수행해야 할 의미적 조정 영역임을 실증적으로 확인하였다.

[표 8] 정량·정성 분석 종합 결과 비교표

[Table 8] Comparison of Comprehensive Quantitative and Qualitative Analysis Results

구분	세부 항목	평가 지표	AI 단독 생성	AI + 디렉션 개입	차이(Δ)	해석
	형태 구조 안정성	SSIM (0~1)	0.58	0.72	+ 0.14	인간 개입 시 캐릭터 외형 안정성 증가
	의미 정합도	CLIP-Sim (I-T)	0.24	0.31	+ 0.07	텍스트-이미지 의미 일치도 향상
	스타일·배경 톤	Mean Color Deviation(ΔE)	14.8	7.2	- 51%	색채·조명 톤 균일성 크게 향상
정성 분석 (전문가)	서사 흐름 자연성	Likert 5점	3.2	4.3	+ 1.1	장면 전환 및 감정 흐름 더 자연적
	캐릭터 정체성 유지	Likert 5점	3.3	4.4	+ 1.1	‘동일 인물성’ 지각 강화
	연출 의도 전달	Likert 5점	3.4	4.2	+ 0.8	컷 구성 및 강조 효과 명료
정성 분석 (독자)	감정 몰입 유지	Likert 5점	3.1	4.1	+ 1.0	감정선 실마리 유지
	해석 명료성	자기보고식	중단 지점 18회	12회	- 32%	장면 해석 단절 감소
	시선 이동 자연성	주관 평가	3.3	4.2	+ 0.9	스크롤 기반 읽기 흐름 안정

4. 결론

본 연구는 생성형 인공지능을 웹툰 제작 과정에 적용할 때 발생하는 시각적 일관성 문제를 중심으로, AI 단독 생성 조건과 인간 디렉션 개입 조건을 비교·분석하였다. 연구 결과, 생성형 AI는 형태적 반복성과 배경 톤 유지와 같이 표면적 시각 요소의 안정화에는 효과적이지만, 감정 흐름의 조절, 장면 전환의 리듬, 서사적 강조와 같은 고차원적 의미 조정 영역에서는 한계를 지남이 확인되었다. 즉, 생성형 AI는 웹툰 제작을 전면적으로 대체하는 장치가 아니라, 시각 작업의 부담을 완화하고 기초 스타일을 유지하는 보조 생산 체계로 이해하는 것이 타당하다. 또한 본 연구는 시각적 일관성을 형태적·서사적·환경적·인지적 네 가지 구성 요소로 구조화하고, 각 요소를 평가하기

위한 정량·정성 복합 분석 지표 체계를 제안하였다. 이를 통해 독자의 서사 경험 속에서 일관성이 어떻게 감지되고 유지되는가라는 체계적인 판단 기준을 제시하였다. 연구 결과는 생성형 AI 기반 웹툰 제작에서 AI와 인간 디렉터의 역할 분담이 필수적임을 시사한다. AI는 스타일 반복, 배경 표현, 톤 유지 등 재현·생성 영역을 담당하고, 인간 디렉터는 감정 연속성, 관계 변화 해석, 장면 강조 지점 결정 등 의미 조정 및 연출적 통제 영역을 수행하는 것이 효율적이다. 이는 향후 웹툰 제작 공정이 단일 창작자 중심 구조에서 AI-디렉터 협업형 제작 체계로 전환될 가능성을 제기한다. 또한, 본 연구는 생성형 인공지능이 웹툰 제작 과정에서 수행할 수 있는 역할과 한계를 규명하고, 시각적 일관성 제어 모델의 필요성을 제시하였으나, 다음과 같은 후속 연구가 요구된다. 첫째, 본 연구는 제한된 장면 단위 시퀀스를 중심으로 분석을 수행하였기에, 장기 연재 환경에서의 일관성 유지 효과를 검증하기 위해 회차 단위·시즌 단위의 확장 연구가 필요하다. 이는 실제 웹툰 산업의 연재 구조를 반영한 시간적 일관성 평가 체계 정립과 연결될 것이다. 둘째, 본 연구에서 설정한 정성 평가 기준은 전문가 그룹과 독자 인지 기반으로 수행되었으나, 시각·인지 행동 추적 기반의 객관적 독해 실험(Eye-tracking, Scroll-trace Analysis)을 추가할 경우, 일관성 판단에 대한 보다 행동 데이터 기반 근거를 확보할 수 있을 것이다. 마지막으로, 생성형 AI의 활용은 단순 기술 문제를 넘어 저작권·작가성·윤리성과 직접 연결된다. 따라서 AI 공동 제작물이 창작 주체성 및 산업 생태계에 미치는 문화·제도·법규적 논의 역시 후속 연구로 병행되어야 한다.

References

- [1] J. H. Kim and Y. Yu, "Platformizing Webtoons: The Impact on Creative and Digital Labor in South Korea," *Social Media + Society*, vol. 5, no. 4, pp. 1-12, Oct. 2019, doi: 10.1177/2056305119880174.
- [2] H. Cho, "The Platformization of Culture: Webtoon Platforms and Media Ecology in Korea and Beyond," *Journal of Asian Studies*, vol. 80, no. 1, pp. 73-93, Feb. 2021, doi: 10.1017/S0021911820001896.
- [3] J. O. Oh, "An Analysis of the Division of Labor in the Webtoon Production Ecosystem," *Journal of Korea Cartoon & Animation Studies*, no. 59, pp. 33-52, June 2020, doi: 10.7230/KOSCAS.2020.59.33.
- [4] L. Zhang, A. Rao, and M. Agrawala, "Adding Conditional Control to Text-to-Image Diffusion Models," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, Oct. 2-6, 2023, pp. 3813-3824, doi: 10.1109/ICCV51070.2023.00355.
- [5] S. McCloud, *Understanding Comics: The Invisible Art*, Harper Collins, 1993.
- [6] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, June 18-24, 2022, pp. 10684-10695, doi: 10.1109/CVPR52688.2022.00350.
- [7] K. Yu, "A Study on Webtoon Generation Using CLIP and Diffusion Models," *Electronics*, vol. 12, no. 18, pp. 3983-3996, Sep. 2023, doi: 10.3390/electronics12183983.

- [8] A. Elgammal, "AI Art and the Role of the Artist," *Leonardo*, vol. 53, no. 4, pp. 424-428, Aug. 2020, doi: 10.1162/LEON_a_01804.
- [9] E. Hu, Y. Shen, P. Wallis, Z. Li, V. Choudhary, and P. Abbeel, "LoRA: Low-Rank Adaptation of Large Language Models," in *International Conference on Learning Representations (ICLR)*, Online, Apr. 25-29, 2022, pp. 1-13.
- [10] N. Cohn, B. Cardoso, B. Klomberg, A. Yang, and M. Orrego, "The Visual Language Research Corpus: An Annotated Corpus of Comics from Asia, Europe, and the United States," *Language Resources and Evaluation*, vol. 57, no. 4, pp. 1729-1744, Dec. 2023, doi: 10.1007/s10579-023-09673-0.