Studies in Media and Communication
Vol. 14, No. 2; June 2026
ISSN: 2325-8071 E-ISSN: 2325-808X
Published by Redfame Publishing
URL: http://smc.redfame.com

Overview of the Application of Generative Artificial Intelligence in Film Production: Algorithms, Tools, and Future Trends

Linze Li[1], M. Asyiek Bin Mat Desa[1], Tiancheng Li[2], Wenjie Li[3]

[1]School of the Arts, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia

[2]Jiangnan University, China. Email: 7210306017@stu.jiangnan.edu.cn

[3]Guangzhou Xinhua University, China. Email: liwenjie_jessic@163.com

Correspondence: M. Asyiek Bin Mat Desa, School of the Arts, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia. Email: asyiek@usm.my

Received: October 27, 2025 Accepted: December 16, 2025 Online Published: December 17, 2025

doi:10.11114/smc.v14i2.8095 URL: https://doi.org/10.11114/smc.v14i2.8095

Abstract

The rapid evolution of generative artificial intelligence (GenAI) is transforming the film industry. This article reviews key GenAI algorithms, including GANs, VAEs, diffusion models, and transformer-based architectures, and explores their application across various stages of film production, from scriptwriting to post-production. Through case studies such as The Frost and Netflix's AI-assisted projects, the study illustrates GenAI's creative potential and workflow innovations. It also addresses critical ethical and legal concerns, including authorship disputes, deepfakes, and algorithmic bias. Finally, the paper outlines future directions, such as multimodal model integration and AI-human co-creation, advocating for a responsible and human-centered implementation of these technologies.

Keywords: generative AI, film production, multimodal models, diffusion models, ethical AI, AI in cinema, creative automation

1. Introduction

Generative AI, especially transformer-based models, has reshaped early filmmaking stages by automating narrative tasks—plotting, dialogue, and character arcs—and redefining the boundary between human creativity and machine authorship in screenwriting, design, and VFX (Anantrasirichai & Bull, 2022).

Historically, the creative practice of filmmaking has involved human sensibility, intuition, and artistic intention. However, the emergence of generative models—namely, GANs, VAEs, Diffusion Models, and Transformer-based systems—challenges the traditional views of authorship and originality (Bengesi et al., 2024). Such algorithms can independently generate scripts, photorealistic images, and original soundtracks and even direct the visual style by learning aesthetic rules(Xu, 2025).The implication is a hybrid model of creativity in which algorithmic tools function as co-authors or idea generators.

Generative Adversarial Networks have proven successful in synthesizing realistic characters, visual style transfer, and building virtual environments. To illustrate how these technical capabilities translate into practice, contemporary film pipelines now employ GAN-based facial animation and deepfake techniques for character design and digital doubles(Mittal et al., 2024). On the other hand, transformer-based models, such as GPT and Stable Diffusion, are multimodal: they produce narrative scripts and visual output, thus further widening the scope of AI-driven storytelling (Bansal et al., 2024; Sinha et al., 2024).

However, the use of generative AI in filmmaking has recently crossed the threshold from playing with technology for its own sake to a business strategy. According to a UnivDatos market report from 2024, generative AI in film will cross $5.5 billion by 2032, its most valuable use case emerging in content generation, virtual actors, and automated dubbing. Due to streaming service demands for new modalities of giving content robust economic viability-scalability, costs minimized in short-form content, particularly for global localization-, these transformations have been prompted. (‘Generative AI In Movies Market Size’, 2024; Yang et al., 2024).

Even the previously resistant auteur cinema has started using GenAI for audience testing, alternative endings, and global localization via voice cloning and automated lip-syncing. MIT Sloan (2023) reported that these practices raise serious legal and ethical questions about authorship transparency and intellectual property (Davenport & Bean, 2023; Y. Huang et al., 2023).

However, while commercial use of GenAI is on the rise, there is an intense debate in scholarly circles about whether GenAI enhances or diminishes human creativity. In this regard, Xu (2025) and Anantrasirichai and Bull (2022) argue for human-AI collaboration to augment imagination, while Garcia (2024) expresses concern that aesthetic homogenization and the loss of narrative authenticity might arise from overreliance on machine-generated content (Anantrasirichai & Bull, 2022; Garcia, 2024).

These debates also find expression in legal and ethical questions. Deepfake and likeness-generation technologies could violate privacy, consent, and fair compensation. Indeed, The Guardian (2024) reports that entire scenes have been created without actors' participation, deepening the ambiguity surrounding digital rights and ownership. The lack of regulatory clarity puts producers and studios in a precarious legal position(Horton, 2024).

While these technical and creative opportunities continue to expand, they also intersect with unresolved ethical and legal concerns, prompting ongoing debates about authorship, transparency, and the responsible use of generative technologies. The substance of productions using AI-generated visuals and dialogue, such as The Frost (Y. Huang et al., 2023; Oberting, 2024), put to test the capabilities of these technologies in producing coherent and attractive narratives with little human input at all (Yang et al., 2024). Such projects serve as testbeds for distilling insights into how GenAI might supplement creative storytelling, with insights gained from them relating to the limitations of the systems involved- their biases in training data, for instance, or their inconsistency in generating long-form narratives(Totlani, 2023).

This interplay between computational logic and artistic intuition represents a paradigmatic shift in the art of cinema. It urges scholars, engineers, and artists to reconsider filmmaking as a non-linear, human-only process but rather as a hybrid, algorithm-enhanced collaboration (Y. Huang et al., 2023; Hutson & Smith, 2025; Santoso & Wijayanti, 2024; Trees, 2024; Uddin et al., 2025). This review intends to map out this evolving terrain by exploring foundational GenAI models, their creative implications, and forecasting future trajectories.

The present study adopts a qualitative literature review approach (Bandara et al., 2015; Onwuegbuzie et al., 2012; Rozas & Klein, 2010; Schryen, 2015). It is based on academic research, industry reports, and technical documentation from the last ten years. The analysis's thematic structuring allows for the classification of GenAI applications in filmmaking across pre-production, production, and post-production stages, thus offering insight into emerging trends, core technologies, and ongoing challenges in AI-integrated filmmaking.

2. Foundations of Generative AI Algorithms

Generative AI has become central to digital content creation, with core models—GANs, VAEs, diffusion models, and hybrids—enabling realistic data synthesis. This section reviews their theoretical foundations and creative applications, particularly in film, as outlined by Sharma et al (Cao et al., 2024; Lin et al., 2025; Ramalakshmi & Asha, 2024; Sharma et al., 2024; Truong et al., 2025).

2.1 Generative Adversarial Networks (GANs)

Goodfellow et al. (2014) proposed the concept of Generative Adversarial Networks, which consists of a generator and a discriminator, and both networks are trained in a minimax game to create samples similar to those in the training set (Goodfellow et al., 2014; Jabbar et al., 2021; Jin et al., 2020).

Gui et al. (2020) have written a comprehensive review of GAN variants such as DCGAN, StyleGAN, CycleGAN, and conditional GANs, which are proposed to enhance the generation quality or control the output attributes. These have been widely adopted in applications such as image synthesis, video frame generation, and artistic style transfer in cinematic production, among others. (Gui et al., 2021). Ghojogh et al. (2021) discuss in detail how different forms of losses have been introduced in GANs, such as Wasserstein loss and feature matching, to stabilize the training and reduce problems like mode collapse(Ghojogh et al., 2021). Chakraborty et al. (2023), in their parallel work, review the decade-long development of GANs, which have transformed several industries, especially those related to creative and entertainment applications. (Chakraborty et al., 2024).

Table 1. Overview of GAN Architectures and Their Enhancements

GAN Variant	Key Feature	Application in Film Production
DCGAN	Deep convolutional structure	Scene synthesis, basic visual effects
CycleGAN	Unpaired image translation	Style transfer, weather/time changes
StyleGAN	Hierarchical latent spaces	Character design, facial animations
cGAN	Conditioning on input labels	Scene-to-scene consistency

Table 1 summarizes key GAN variants and their primary film production applications. Notably, StyleGAN excels in generating high-resolution, human-like faces, facilitating virtual casting and digital actor creation.

2.2 Variational Autoencoders (VAEs)

Variational Autoencoders, first introduced by Kingma and Welling in 2014, are a probabilistic approach to data generation. (Kingma & Welling, 2013). Unlike GANs, VAEs aim to learn data distributions by encoding inputs into latent space and reconstructing them, optimized via a loss combining reconstruction error and Kullback-Leibler divergence for regularization (Asperti & Trentin, 2020; Shao et al., 2020; Song et al., 2019).

This makes the VAE particularly suitable for smoothly learning interpretable latent representations, which can be applied in film for scene generation, facial morphing, interpolative transitions between frames, etc. Ghojogh et al. (2021) compare the VAE with the GAN and add that though the VAE sometimes delivers blurrier results, its training is much more stable and interpretable (Ghojogh et al., 2021). Additionally, it is widely used for generative data augmentation, wherein it improves a film analytics-based downstream task training set, Chen, Yan, and Zhu (2024) (Chen, Yan, et al., 2024).

2.3 Diffusion Models

Diffusion models represent a new frontier in GenAI. Sohl-Dickstein et al. (2015) define a forward process that adds noise to data and a reverse process that denoising it to reconstruct the original input(Sohl-Dickstein et al., 2015). This allows highly stable training and has generated exceptionally high-resolution and coherent images.

Joshi (2025) and Rombach et al. (2022) elaborate on Latent Diffusion Models (LDMs), wherein the diffusion process takes place in a compressed latent space instead of pixel space, thus vastly improving the computational efficiency of the diffusion models discussed here. They have been used to generate high-definition background information, several variations of lighting conditions fit for film, and even an artistic emulation such as post-production(Joshi, 2025; Rombach et al., 2022).

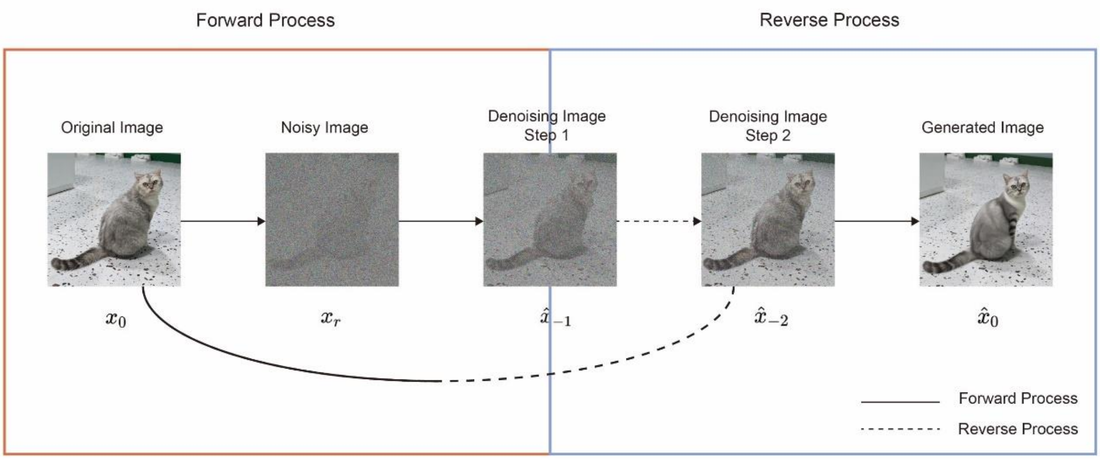

Figure 1. Diffusion Process Visualization

This flowchart illustrates the two-phase structure of diffusion models:

Forward Process: Data + Noise → Latent Representation

Reverse Process: Latent Representation → Clean Data

Such visualization aids in understanding how gradual denoising reconstructs complex patterns, enabling realistic video frames from pure noise.

2.4 Adversarial Autoencoders and Hybrid Models

To combine the strengths of VAEs and GANs, adversarial autoencoders (AAEs) were developed. These models use an encoder-decoder architecture similar to VAEs but incorporate an adversarial training mechanism to match the aggregated posterior of the latent space with a prior distribution.

Ghojogh et al. (2021) discuss how AAEs facilitate better latent space modeling and sharper output generation compared to traditional VAEs (Ghojogh et al., 2021). These models are particularly suitable for character animation pipelines where fine control over latent variables is required.

Table 2. Comparative Table of Generative Models

Model Type	Core Mechanism	Pros	Cons
GAN	Adversarial training	High-quality, realistic outputs	Training instability, mode collapse
VAE	Variational inference	Stable training, latent structure	Lower output resolution
Diffusion	Noise reversal	Coherent, high-res outputs	Slower generation time
AAE	Hybrid (VAE + GAN)	Sharp images + structured latent	Complex to train

This comparison aids in selecting appropriate algorithms based on production needs, such as speed, resolution, or controllability.

2.5 Mathematical Underpinnings and Optimization

Understanding the mathematical formulation of these models is essential for their implementation in creative workflows. In GANs, the min-max optimization problem can be written as:

$$min_G max_D V(D,G) = \mathbb{E}_{x\sim p_{data}}[log D(x)] + \mathbb{E}_{z\sim p_z}\left[log\left(1 - D\big(G(z)\big)\right)\right]$$

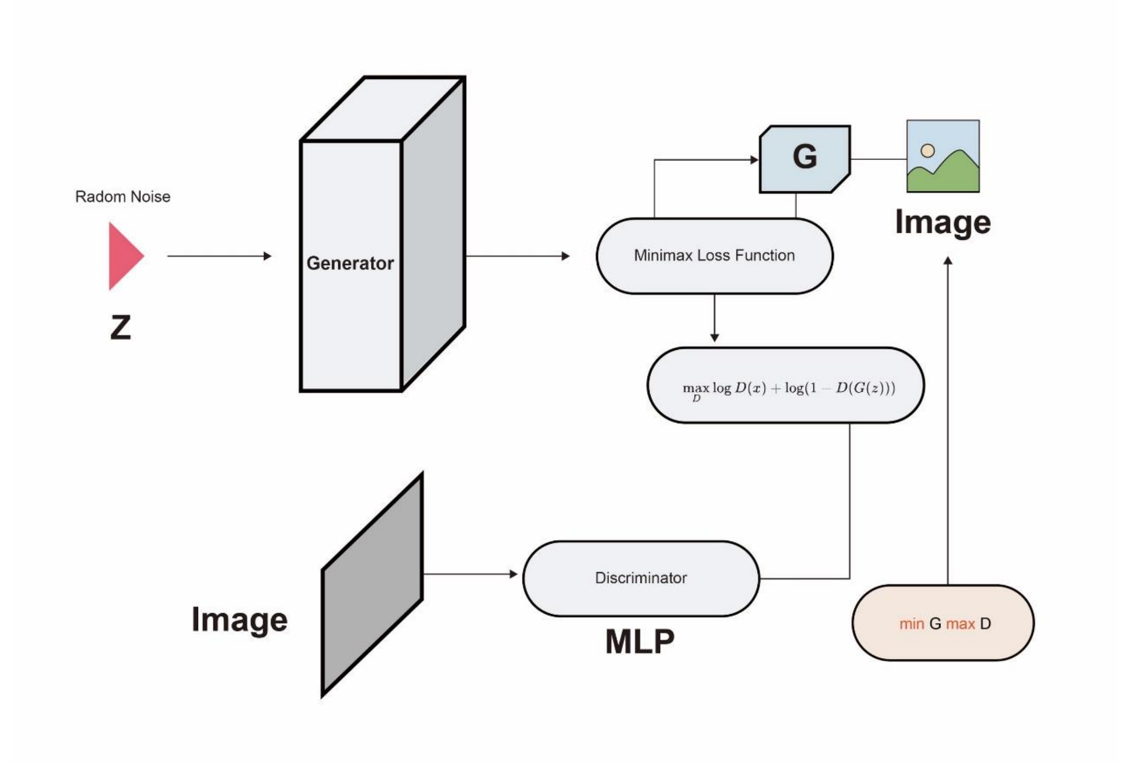

Figure 2. GAN Loss Derivation Flowchart

This flowchart shows the process of GAN training with its loss formulation:

Sample Real Data $x \sim p_{data}$

Sample Latent Vector $z \sim p_z$

Generate Fake Data $G(z)$

Compute Discriminator Outputs $D(x)$, $D\big(G(z)\big)$

Discriminator Loss: $\log D(x) + \log\left(1 - D\big(G(z)\big)\right)$

Generate Loss: $\log\left(1 - D\big(G(z)\big)\right)$ or $-\log\left(1 - D\big(G(z)\big)\right)$ in non-saturating form

Optimize using gradient descent on respective losses

This schematic illustrates how GANs balance their two networks during adversarial training and highlights where each component of the loss function fits within the iterative training loop.

For VAEs, the evidence lower bound (ELBO) is:

$$\mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(x|z)}[\log p_\theta(x|z)] - D_{KL}\left(q_\phi(x|z) \parallel p(z)\right)$$

Diffusion models use a Markov chain framework with a Gaussian noise schedule. Training involves estimating the noise added at each step and minimizing the mean squared error between true and predicted noise.

Generative AI models—GANs, VAEs, and diffusion architectures—enhance visual creativity across film production stages. Each suits specific tasks, from rapid prototyping to photorealistic rendering. Their integration boosts productivity, empowers digital artists, and enables new aesthetics, while raising platform, tooling, and ethical considerations in practical deployment.

The following paragraphs explore how these can be integrated into production, including deployment at the platform level, tooling ecosystems, and ethical considerations.

3. Applications of Generative AI in Film Production

Generative AI (GenAI) is transforming the landscape of world cinema by automating creative tasks and improving efficiencies. Building upon the algorithms discussed in the previous section, this part examines how these models function within actual production environments, demonstrating how technical mechanisms translate into real-world creative workflows.

3.1 Pre-Production: Scriptwriting, Storyboarding, and Concept Design

Recently, GenAI has become a significant resource in pre-production, particularly in scriptwriting, character design, and visual ideation. For instance, LLM-based tools like ChatGPT enable fast dialogue writing, scene descriptions, and alternative plot suggestions. Mozelius and Humble (2024) noted that generative models can dramatically reduce ideation time, especially for formulaic and genre-specific content(Mozelius & Humble, 2024).

Similarly, text-to-image models like MidJourney and DALL·E allow production designers to quickly visualize concepts for sets, costumes, and entire scenes with simple text prompts. These are, of course, only starting points for the concept art guiding the rest of the creative team. Indeed, Rombach et al. (2022) showed that latent diffusion models could be more scalable and visually appealing than traditional 2D sketching tools(Rombach et al., 2022).

Table 3. Comparison of Typical AI Tools in Film Production

Tool	Functionality	Advantages	Disadvantages	Commercialization Level
ChatGPT	Scriptwriting, dialogue, idea generation	Fast iteration, low cost, customizable prompts	Generic output, lacks deep narrative logic	High
MidJourney	Concept art generation	Stylized output, detailed control, visual aesthetics	Limited realism, requires prompt engineering	Medium
Runway ML	Video editing, motion tracking, VFX	User-friendly, real-time rendering, diverse effects	Limited by model training scope	High
Synthesia	AI-generated avatars for narration	Multilingual support, scalable video narration	Uncanny valley issues, limited emotion range	Medium
D-ID	Face animation from static images	Realistic facial movement, fast processing	Expression limitations, requires clean input	Medium

This table summarizes how various AI tools perform in terms of function, usability, and deployment within a typical production cycle.

3.2 Production: Set Design, Virtual Filming, and Synthetic Actors

Generative models support several innovative applications during the active filming stage. In virtual production, platforms such as Unreal Engine, integrated with GenAI models, allow directors to preview fully realized environments before shooting (Al Naqbi et al., 2024). Diffusion-based systems generate photorealistic synthetic settings, which can be projected onto LED walls to construct immersive, responsive digital stages (Chamola et al., 2024; Hutson, 2024; Leininger et al., 2025; Viktoriia, 2024).

These technologies also enable the creation of "digital doubles" through AI-generated avatars and deepfake techniques.

Though ethically contentious, such methods have been experimentally employed to complete scenes involving absent or deceased actors (Cohen & Giryes, 2022; Hutson, 2024),illustrating both the creative potential and ethical complexity of GenAI in actor representation.

This radar chart visualizes how commonly used AI tools—ChatGPT, MidJourney, Runway ML, Synthesia, and D-ID—are mapped onto different production phases:

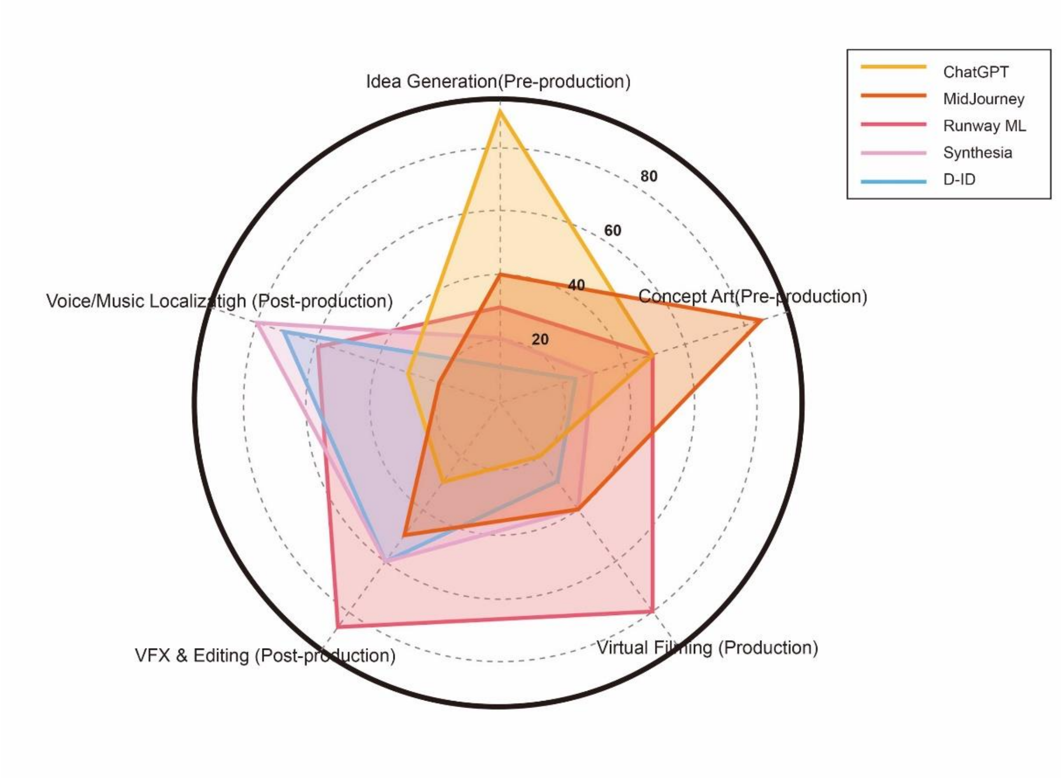

Figure 3. Radar Chart of AI Tool Applicability Across Film Production Stages

ChatGPT: strong in pre-production (90%), weak in post (30%)

MidJourney: high in concept art (85%), low in final production (40%)

Runway ML: robust in production/post-production (80%-85%)

Synthesia/D-ID: medium in post (60%), low in conceptual stages (20%-30%)

These distributions reveal how AI capabilities are currently more mature in pre-visualization and post-production than in live filming.

3.3 Post-Production: Editing, VFX, Audio and Localization

Most post-production processes have benefited from GenAI. Tools such as Runway ML enable editors to do VFX, bg removal, and scene transitions with minimal manual keyframing. Joshi (2025) claims that these tools reduce rendering times and the cognitive load on editors and guide attention better toward narrative coherence (Epstein et al., 2023; Hutson & Smith, 2025; Joshi, 2025; Somaini, 2023).

In audio processing, AI-based voice synthesis tools allow dubbed voices to match mouth movements and enable real-time multilingual localization (Pataranutaporn et al., 2021). For example, Synthesia provides digital voice clones of actors, making logistics for the international release smoother and cheaper.

Generative AI also helps in composing music and tracking the mood in scenes. Tools akin to MidJourney and MuseNet can generate soundtracks according to the appropriate visual and emotional cues (Dong, 2024; He et al., 2025; Kang et al., 2024; Mitra & Zualkernan, 2025; H. Park et al., 2024; Zhong et al., 2024).

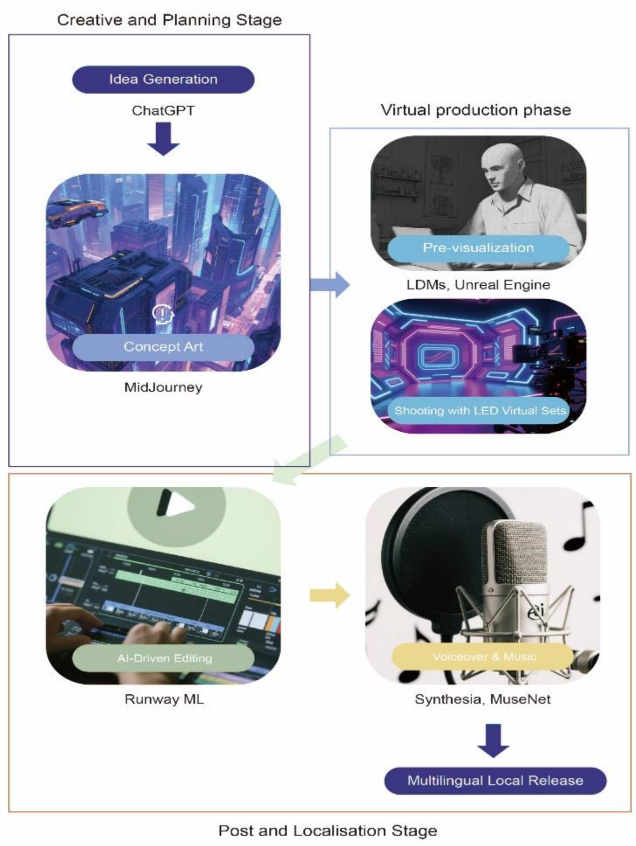

Figure 4. Flowchart of AI Technology Integration in Film Workflow

1→Idea Generation (ChatGPT) → 2. Concept Art (MidJourney) → 3. Pre-visualization (LDMs, Unreal Engine) → 4. Shooting with LED Virtual Sets → 5. AI-Driven Editing (Runway ML) → 6. Voiceover & Music (Synthesia, MuseNet) → 7. Multilingual Local Release

This flowchart illustrates how tools build upon one another, contributing to a more cohesive and semi-automated creative pipeline.

3.4 Industry Applications and Case Studies

Both large studios and independent creators are integrating GenAI into their workflows. For example, Netflix is said to be using GPT-based models for auto-generating subtitles and writing trailer scripts (Balestri et al., 2024; J. Huang et al., 2025; Solanki & Khublani, 2024). Similarly, Sony Pictures tested synthetic actors generated with StyleGAN for short film production (Cao et al., 2023).

Interest in GenAI among academics is not lagging. Mozelius and Humble (2024) observed that film students using GenAI tools demonstrated faster turnaround times and increased creative confidence. Ren et al. (2023) found that diffusion-based systems outperform traditional animation tools in facial tracking and physics-based rendering (Ren et al., 2025).

Cohen and Giryes (2022) emphasize GANs' effectiveness in reconstructing missing frames for film restoration. In a low-resource context like independent cinema, Karras et al. (2020) outline possible paths for GAN training with limited data (Cohen & Giryes, 2022; Karras et al., 2020).

3.5 Limitations and Ethical Implications

While GenAI demonstrates significant capabilities, its integration comes with several challenges. This includes the risk of over-relying on pre-trained models, leading to the homogenization of artistic outputs. Moreover, deepfake technology raises serious ethical concerns regarding consent, identity, and misrepresentation.(Chapagain et al., 2024).

On the legal side, copyright law for AI-generated works remains unsettled, especially for proprietary materials used in model training. According to Geiger (2024), the intellectual property paradigm needs to be updated to balance the interests of rights holders and algorithmic innovation (Geiger, 2024).

The GenAI outputs can be hit and miss. As Ni et al. (2023) elaborated, generative models have yet to master the logical coherence of extended sequences or contextual ambiguity across different modalities (Ni et al., 2024).

With the development of model architecture and contextual reasoning, the focus of the development of generative artificial intelligence must be on long-term ethical, legal, and artistic impacts. In this regard, key directions for the future include model transparency, content attribution standards, and a fair framework created jointly by humans and artificial intelligence.

4. Case Studies and Real-World Applications

4.1 Representative Cases

The use of generative artificial intelligence in film production has progressed from speculative tests to applied case studies with demonstrable impacts. Such applied use cases not only testify to the technical capabilities of GenAI but also reshape production workflows, economic structures, and aesthetic conventions. Representative examples range from entirely AI-generated short films like The Frost through scriptwriting pilots to micro-level analyses of creator workflows.

4.1.1 The Frost: A Short Film Produced Entirely with AI Assistance

“The Frost” is a paradigm-defining example of end-to-end AI filmmaking (The Frost – Waymark, 2024). It used tools such as Runway ML, MidJourney, and ChatGPT across all production stages, including storyboarding, dialogue generation, visual composition, and voice synthesis. It was identified by Georgia State University as a paradigmatic case of full-stack GenAI implementation (Georgia State University, 2025).

Research into the film has shown that The Frost evidences significant time and cost savings in production. For example, character design, location scouting, and visual referencing—all processes that take up to 72 hours—were expedited using the MidJourney platform (Evans, 2024; Hemraj, 2025; MILLER, 2024). Text-to-video tools similarly compress the animation and rendering stages of post-production, which are traditionally very labor-intensive.

Table 4. Comparison of AI Usage in Real Movie Projects

Project	AI Tools Used	Stage(s) of Use	Cost Reduction	Time Saved	Feedback on Quality
The Frost	ChatGPT, MidJourney, Runway ML	Pre-production, Post-production	~60%	~70%	High visual fidelity, limited motion realism
Fender AI Commercial	Stable Diffusion, ElevenLabs	Pre-production, Voice-overs	~40%	~50%	Smooth integration, needs human refinement
Sophia-in-Audition	Custom LLM, UltraStage	Virtual Acting, Real-time Rendering	N/A	N/A	Breakthrough in robotic performance
Netflix Script Project	GPT-4, in-house generators	Scriptwriting, Character Dialogue	N/A	~30%	Mixed; human writers needed to revise
Project	AI Tools Used	Stage(s) of Use	Cost Reduction	Time Saved	Feedback on Quality

The table underscores the transformative impact of AI across different phases of production. While cost and time savings are significant, quality feedback remains mixed, necessitating hybrid models combining AI and human expertise.

4.1.2 Netflix and OpenAI Collaboration: Script Experimentation

Netflix has rumoredly done some pilot tests in partnership with OpenAI, attempting to explore script co-writing for specific genres of series and pilot concepts. However, this is confidential and cannot be discussed publicly(Chiappani, 2024; Townsend, 2024). U.S. entertainment industry insiders told third parties that AI tools were used to produce first drafts of dialogue trees, scenes, and character arcs. Actual people then wrote these initial drafts, who went on to finish the story in a way that maintained narrative coherence and emotional truth. The Github teddy below would likely need to be shown for reference; this, too, may be impossible to discuss openly (Mirowski et al., 2023).

The experiment, still in its pilot stage, suggests a crucial change in the dynamics within writers' rooms. Instead of making human creativity redundant, it possibly serves as an accelerant for ideas. Indeed (Chandrasekera et al., 2025; Geyer & Rosignoli, 2024; Zhou & Lee, 2024), Haase (2023) states that AI helps screenwriters break their creative blocks by generating multiple plot continuations within minutes (Haase & Hanel, 2023). However, concerns about originality and ownership are formidable (Oshiesh, 2025).

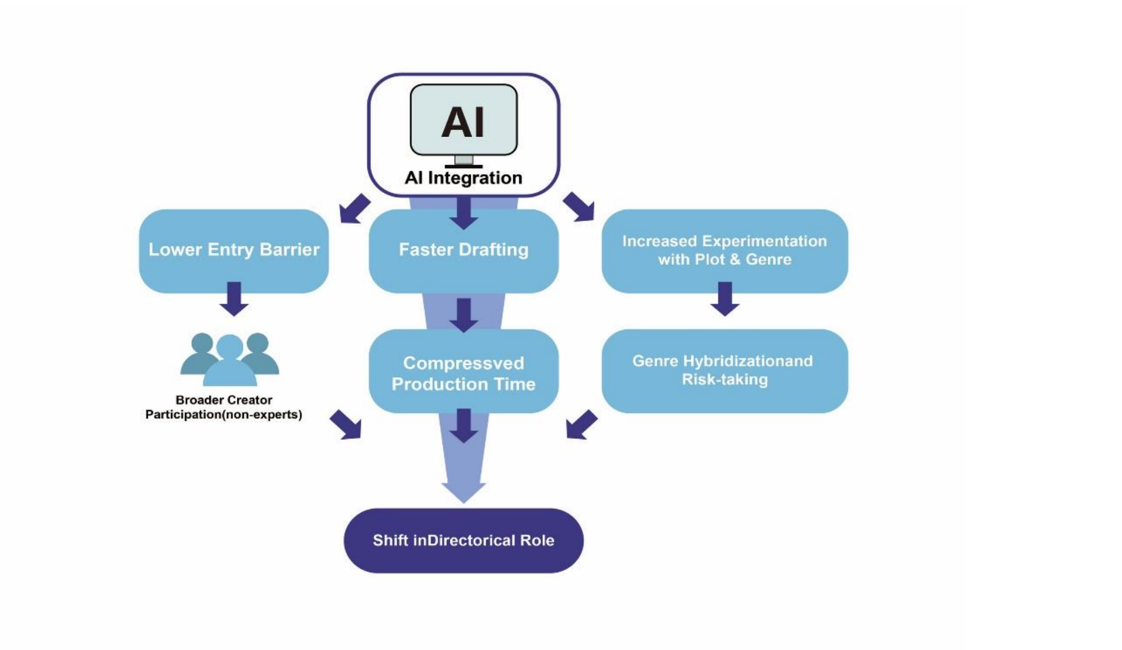

Figure 5. Causal Flowchart of AI Tools on Changes in Filmmaker Behavior

1→AI Integration → 2. Reduced Technical Barriers → 3. Faster Iterations → 4. Higher Volume of Creative Prototypes → 5. Shift in Human Role from Execution to Supervision → 6. Rethinking of Authorship and Copyright

This causal flow highlights the systemic changes in film production environments. The introduction of GenAI reduces the need for deep technical expertise in early production stages, allowing filmmakers to focus more on ideation and refinement.

Recent research highlights creators' adaptation to GenAI via modular workflows. Independent filmmakers use tools like MidJourney for concept art, ChatGPT for synopses, and Synthesia for animatics (Anantrasirichai et al., 2025; Laco, 2024).

These enhance creative control and feedback cycles. Netflix's experiments show AI reshaping roles and workflows, yet ethical norms, training methods, and collaboration models remain unresolved (Ching & Mothi, 2025; Przegalinska & Triantoro, 2024; Serena, 2025).

5. Challenges and Ethical Considerations

5.1 Legal and Copyright Issues

The rapid development of generative AI has caused profound challenges to legal interpretation, especially in copyright. Indeed, as Lee (2024) notes, U.S. copyright law presently excludes works produced entirely by machines from eligibility for protection unless a human author can be identified (Gray, 2023; Lee, 2024). In that light, Thaler v. Perlmutter reaffirmed the principle that copyright eligibility requires human authorship (Mainini, 2023; Rezek, 2024; Rosati, 2025; Seshadri, 2021). That creates a problem for filmmakers who may wish to use AI tools in scripting, editing, and other aspects of production.

While a few jurisdictions have moved toward limited recognition of AI-generated content, there is no international consensus. For example, the Copyright, Designs and Patents Act 1988 of the United Kingdom states that attribution is to be given to the person who "makes the arrangements necessary for the creation of the work" (Varghese, 2023). However, such approaches are not uniformly adopted, complicating enforcement in cross-border collaborations involving AI.

The synthetic replication of human likeness by AI raises significant ethical and legal concerns. As Verma (2024) notes, unauthorized use can violate privacy and publicity rights. Cases like Bruce Willis' likeness misuse and laws such as Tennessee's ELVIS Act highlight the urgent need for regulatory reform (Hendrickson, 2025; Verma, 2024).These efforts underscore the need for updated legal frameworks capable of managing new threats posed by GenAI tools.

5.2 Tension Between Creativity and Technology

5.2.1 Will Human Creativity Be Marginalized by Technology?

In their increasingly important role in creative processes, AI tools have raised doubts regarding the actualization of creativity in people. Epstein et al. (2023) suggest that AI might produce aesthetically appealing yet devoid of profound emotion. Thus, more excellent reliance on AI risks culturally making our creations unmarked and uniform (Hughes et al., 2021).

On the other hand, others see AI as something that can support and supplement human creativity instead of supplanting it. Townsend (2024) states that GenAI does not replace human imagination and creativity but has the potential to be a "creative assistant" to overcome writer's block and visualize an entire scene faster if it is used responsibly (Townsend, 2024; Washington, 2023). This perspective views AI not as a threat, but as a tool that expands the creator's palette.

5.2.2 The Impact of Template-Based Creation on Artistic Quality

Template-based AI creation may also threaten diversification in an artistic context. According to Caramiaux (2025), AI-powered tools create content mainly based on their training data, which can generate repetitive and derivative content. This threatens the originality and diversity underpinning higher artistic values. This may again lead to less innovation (Caramiaux et al., 2025).

GenAI's democratization of media production through reduced barriers to access may come at the cost of questioning the authenticity and worth of such creations from a commercial standpoint. Indeed, framed by a traditional perspective of artistic merit and originality, permitting AI in creative authorship and ownership may be damaging (Bosch, 2024; Garcia, 2024; Y. Wang, 2023).

5.3 Model Transparency and Bias

Hamilton(2020) proposed that artificial intelligence systems often replicate biases in training data; For example, the darker the skin color, the higher the facial recognition error rate (Hamilton, 2020). Addressing such ethical risks in generative AI requires fairness-aware design, transparent auditing, and alignment between algorithmic processes and social accountability (Arrieta et al., 2020).

A significant concern in GenAI is its "black box" nature—users often cannot interpret how outputs are produced (Arrieta et al., 2020). In film, this opacity risks misalignment with ethical intent. Explainable AI seeks to ensure interpretability, fostering accountability and trust in creative applications (Balasubramaniam et al., 2023; Chamola et al., 2023; Eschenbach, 2021; Hassija et al., 2024; Luk et al., 2024; Saeed & Omlin, 2023; Thalpage, 2023).

6. Future Directions and Emerging Trends

6.1 Multimodal Model Fusion

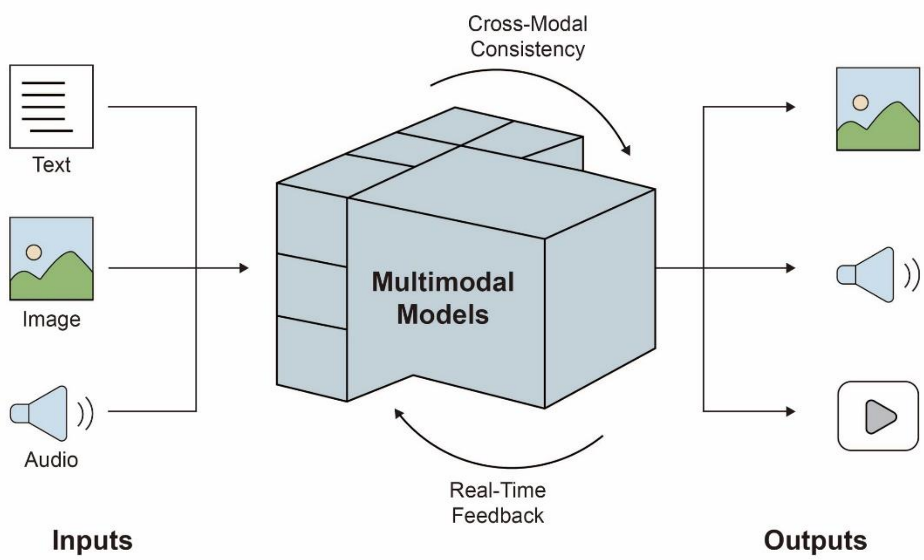

Figure 6. Schematic Diagram of Multimodal Generation Ecosystem Structure

The figure depicts integrating language, image, and audio models into a single generative pipeline supporting prompt-based storytelling. Feedback loops iteratively refine tone, visuals, and sound, enabling real-time, low-latency production workflows in next-generation film studios.

6.1.1 The Integrated Evolution of Text, Image, and Video Generation

Recent efforts in generative AI have concentrated on multimodal data synthesis: a single architecture for text, image, and video generation. This is a paradigm shift compared to the previous generation of unimodal models, and research in cross-modal semantics encoding is oriented toward making plausible multimedia generation from a single prompt (Gao et al., n.d.; Liang et al., 2024; Liu et al., 2024).

Examples of nascent models in this area include Video-LLaMA, indicative of the progressive alignment of linguistic and visual representations. Annotated datasets and a transformer-based backbone, joint embedding, and joint fusion mechanisms could lead to text-driven scene synthesis, where a script defines characters' actions, scene composition, and camera movement (Ataallah et al., 2024; D. Huang et al., 2024; Saeed & Omlin, 2023; Zhang et al., 2023).

In addition, Chen et al. (2024) analyze the performance of a unified generative model for image captioning, visual storytelling, and cinematic scene synthesis tasks. This explores the scalability of such models concerning real-time film concept art design, pre-visualization, and scene simulation (Abootorabi et al., 2025; Chen, Rao, et al., 2024; Yin, n.d.).

6.1.2 Synchronous Generation Models for Audio and Video

Synchronous audio and video generation remain a challenging but crucial aim in GenAI-based filmmaking. Separately coded methods inevitably incur the risk of temporal misalignment and ill-perceptive quality degradation. To mitigate this effect, Wang et al. (2024) present AV-DiT, a transformer-based diffusion framework for audio-visual joint generation. This model, through synchronized dual attention streams over video and audio representations, advances cross-modal coherence and temporal consistency in a primary perceptual quality decodement process (J. Park et al., n.d.; K. Wang et al., 2024).

The diffusion transformer model proposed by Yu et al. (2025) and Xing et al. (2024) has hierarchical scheduling under audio-visual rhythm and semantic preservation. It is a state-of-the-art model in short film trailer generation, embedding a recorded voice, ambient sound, and score. Indeed, the cross-attention map has been proven effective in visualizing how the dialogue cue or ambient text influences the pathway of video generation (Xing et al., 2024; Yu et al., 2025).

These systems represent an emerging paradigm where narrative and aural elements are generated simultaneously, offering end-to-end solutions for AI-driven cinematic production pipelines.

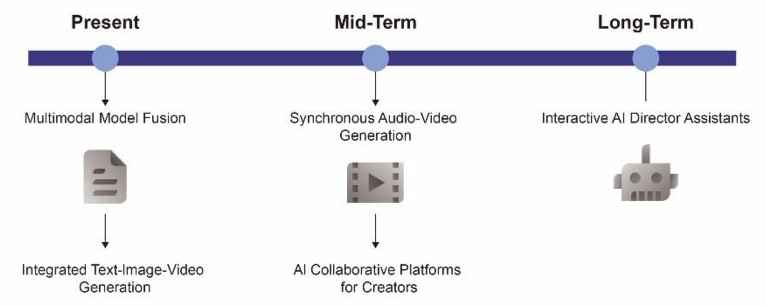

Figure 7. Roadmap of Future Generative AI Development

The roadmap illustrates the evolution of generative AI, from foundational models to real-time multimodal collaboration. Further development could see emotionally adaptive systems directing films under thematic and audience constraints. Such a trajectory emphasizes the pressing need to co-develop ethical frameworks with the technical development itself.

6.2 Intelligent Collaboration and Creative Partner AI

Phillips et al. (2024) described how collaborative artificial intelligence platforms are reshaping film creation by integrating writing, animation, and sound generation tools(Phillips et al., 2024).

This reflects the mature model of human-in-the-loop creativity, in which dynamic user interactions with the AI's suggestions, accepting modifications, and iteratively generating content by reinforcement-learning-based preference tuning result in better cortical coherence and user satisfaction, as Kumar et al. (2024) indeed noted, in static-generation models (Kumar et al., 2024; Mathewson, 2019; Santoso & Wijayanti, 2024).

AI in the role of "director's assistant" has moved from speculative design to a working, functional prototype. The range of interactive AI systems currently being created goes well beyond automation of content orchestration and emotional tone calibration to include everything required to monitor visual continuity throughout a film. For example, Swarnakar (2024) report an agent-based assistant capable of analyzing screenplay sentiment, suggesting blocking patterns(Swarnakar, 2024).

These will use RLHF and instruction-tuned large models to steer their creative suggestions toward a predetermined directorial style. Based on a three-layer control structure of a task parser, a creative engine, and an execution scheduler, Ni et al. (2024) proposed an architecture that positions them as fully-fledged collaborators in next-generation filmmaking(Ni et al., 2024).

7. Conclusion

Generative AI has shifted from an assistive tool to a creative collaborator in film. This review outlines its technical foundations, applications, and ethical challenges. While enabling faster, cheaper, and more expressive workflows, its future depends on aligning innovation with ethical oversight to amplify—rather than replace—human creativity.

Essential paths of future research will include:

Algorithmic transparency transparency requires the development of explainable GenAI systems when high-stakes narrative and visual media creative decisions are at stake.Ethical governance frameworks can only be co-designed with industry stakeholders to assure fair credit attribution, mitigate deepfake misuse, and protect actor likeness rights.Significant input from computer scientists, filmmakers, and legal and ethical academics will help balance discourse on innovation with notions of responsibility.Empirical study of the cognitive, emotional, and economic impacts of these tools on creators will help build respectful systems of professional needs and values.

GenAI is both an opportunity and a burden. Its potential for rich, complex, adaptive content will continue to grow. Its success in film production and human, equitable, ethical, and creative integration into storytelling determines its value. Therein lies the future of cinema: algorithm-enhanced yet human-inspired.

Acknowledgments

Thank you to every author for their contributions to the writing of the paper, as well as to colleagues who assisted in conducting research or commenting on the manuscript. Thank you to the personnel who regularly participate in reviewing and accepting manuscripts, such as peer reviewers or editors, associate editors, and consulting editors of journals where articles are published

Authors contributions

Dr. Li Linze and Dr. Asyiek are responsible for research design and revision. Dr. Li Tiancheng is responsible for data collection. Dr. Li Linze drafted the manuscript, and Dr. Li Wenjie proofread and organized it. Finally, all authors read and approved the final manuscript.

Funding

This study received no financial support

Competing interests

The author declares that this study has no known competitive financial interests or personal relationships and will not affect the work reported in this article.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

Copyrights

References

Abootorabi, M. M., Ghahroodi, O., Zahraei, P. S., Behzadasl, H., Mirrokni, A., Salimipanah, M., ... & Asgari, E. (2025). Generative AI for Character Animation: A Comprehensive Survey of Techniques, Applications, and Future Directions. arXiv preprint arXiv:2504.19056.

Al Naqbi, H., Bahroun, Z., & Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. *Sustainability, 16*(3), 1166. https://doi.org/10.3390/su16031166

Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial intelligence review, 55*(1), 589-656. https://doi.org/10.1007/s10462-021-10039-7

Anantrasirichai, N., Zhang, F., & Bull, D. (2025). Artificial Intelligence in Creative Industries: Advances Prior to 2025. arXiv preprint arXiv:2501.02725.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion, 58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Asperti, A., & Trentin, M. (2020). Balancing reconstruction error and kullback-leibler divergence in variational autoencoders. *Ieee Access, 8,* 199440-199448. https://doi.org/10.1109/ACCESS.2020.3034828

Ataallah, K., Shen, X., Abdelrahman, E., Sleiman, E., Zhu, D., Ding, J., & Elhoseiny, M. (2024). Minigpt4-video: Advancing multimodal llms for video understanding with interleaved visual-textual tokens. arXiv preprint arXiv:2404.03413.

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology, 159,* 107197. https://doi.org/10.1016/j.infsof.2023.107197

Balestri, R., Cascarano, P., Degli Esposti, M., & Pescatore, G. (2024, September). An Automatic Deep Learning Approach for Trailer Generation through Large Language Models. In *2024 9th International Conference on Frontiers of Signal Processing* (ICFSP) (pp. 93-100). IEEE. https://doi.org/10.1109/ICFSP62546.2024.10785516

Bandara, W., Furtmueller, E., Gorbacheva, E., Miskon, S., & Beekhuyzen, J. (2015). Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support. *Communications of the Association for Information systems, 37*(1), 8. https://doi.org/10.17705/1CAIS.03708

Bansal, G., Nawal, A., Chamola, V., & Herencsar, N. (2024). Revolutionizing visuals: the role of generative AI in modern image generation. *ACM Transactions on Multimedia Computing, Communications and Applications, 20*(11), 1-22. https://doi.org/10.1145/3689641

Bengesi, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2024). Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers. IEEe Access. https://doi.org/10.1109/ACCESS.2024.3397775

Bosch, M. J. (2024). *Generation Z's Perceived Authenticity and Engagement with AI-Generated vs. Human-Made Artistic Media Content: The Roles of Content Labeling and Attitude toward Regulation* (Bachelor's thesis, University of Twente).

Cao, H., Tan, C., Gao, Z., Xu, Y., Chen, G., Heng, P. A., & Li, S. Z. (2024). A survey on generative diffusion models. *IEEE Transactions on Knowledge and Data Engineering.* https://doi.org/10.1109/TKDE.2024.3361474

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. arXiv preprint arXiv:2303.04226.

Caramiaux, B., Crawford, K., Liao, Q. V., Ramos, G., & Williams, J. (2025). Generative AI and Creative Work: Narratives, Values, and Impacts. arXiv preprint arXiv:2502.03940.

Chakraborty, T., KS, U. R., Naik, S. M., Panja, M., & Manvitha, B. (2024). Ten years of generative adversarial nets (GANs): a survey of the state-of-the-art. *Machine Learning: Science and Technology, 5*(1), 011001. https://doi.org/10.1088/2632-2153/ad1f77

Chamola, V., Bansal, G., Das, T. K., Hassija, V., Sai, S., Wang, J., ... & Niyato, D. (2024). Beyond reality: The pivotal role of generative ai in the metaverse. *IEEE Internet of Things Magazine, 7*(4), 126-135. https://doi.org/10.1109/IOTM.001.2300174

Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D., & Sikdar, B. (2023). A review of trustworthy and explainable artificial intelligence (xai). *IEEe Access, 11,* 78994-79015. https://doi.org/10.1109/ACCESS.2023.3294569

Chandrasekera, T., Hosseini, Z., & Perera, U. (2025). Can artificial intelligence support creativity in early design processes?. *International journal of architectural computing, 23*(1), 122-136. https://doi.org/10.1177/14780771241254637

Chapagain, D., Kshetri, N., & Aryal, B. (2024, July). Deepfake Disasters: A Comprehensive Review of Technology, Ethical Concerns, Countermeasures, and Societal Implications. In *2024 International Conference on Emerging Trends in Networks and Computer Communications* (ETNCC) (pp. 1-9). IEEE. https://doi.org/10.1109/ETNCC63262.2024.10767452

Chen, Y., Rao, A., Jiang, X., Xiao, S., Ma, R., Wang, Z., ... & Dai, B. (2024). CinePreGen: Camera Controllable Video Previsualization via Engine-powered Diffusion. arXiv preprint arXiv:2408.17424.

Chen, Y., Yan, Z., & Zhu, Y. (2024). *A comprehensive survey for generative data augmentation.* Neurocomputing, 128167. https://doi.org/10.1016/j.neucom.2024.128167

Chiappani, L. (2024). AI in the film industry | Opportunities and challenges. Ki-Company.ai. https://www.ki-company.ai/en/blog-beitraege/artificial-intelligence-in-the-film-industry-opportunities-and-challenges

Ching, V., & Mothi, D. (2025). *AI for Creatives: Unlocking Expressive Digital Potential.* CRC Press. https://doi.org/10.1201/9781003533016

Cohen, G., & Giryes, R. (2022). Generative adversarial networks. arXiv preprint arXiv:2203.00667. https://arxiv.org/abs/2203.00667

Davenport, T., & Bean, R. (2023, June 19). The Impact of Generative AI on Hollywood and Entertainment. MIT Sloan Management Review; Massachusetts Institute of Technology. https://sloanreview.mit.edu/article/the-impact-of-generative-ai-on-hollywood-and-entertainment/

Dong, H. W. (2024). *Generative AI for Music and Audio.* University of California, San Diego.

Epstein, Z., Hertzmann, A., Investigators of Human Creativity, Akten, M., Farid, H., Fjeld, J., ... & Smith, A. (2023). Art and the science of generative AI. *Science, 380*(6650), 1110-1111. https://doi.org/10.1126/science.adh4451

Evans, A. K. (2024). *The Conglomeration of Everything: Shifting Animation Production Practices due to Artificial Intelligence* (Doctoral dissertation, Doctoral dissertation, San Francisco State University).

Gao, F., Liu, M., & Zhou, Y. Cross-Modal Generation of Visual and Auditory Content: A Survey. In *Artificial Intelligence for Art Creation and Understanding* (pp. 84-112). CRC Press. https://doi.org/10.1201/9781003406273-4

Garcia, M. B. (2024). The paradox of artificial creativity: Challenges and opportunities of generative AI artistry. *Creativity Research Journal,* 1-14. https://doi.org/10.1080/10400419.2024.2354622

Geiger, C. (2024). Elaborating a Human Rights-Friendly Copyright Framework for Generative AI. *IIC-International Review of Intellectual Property and Competition Law, 55*(7), 1129-1165. https://doi.org/10.1007/s40319-024-01481-5

Generative AI In Movies Market Size, Growth Report, 2032. (2024). UnivDatos. https://univdatos.com/reports/generative-ai-in-movies-market

Georgia State University. (2025). Lights, camera, algorithm: How AI is transforming filmmaking. https://provost.gsu.edu/2025/04/01/lights-camera-algorithm-how-ai-is-transforming-filmmaking/

Geyer, R., & Rosignoli, A. (2024). The Influence of Generative AI on Creativity in the Front End of Innovation.

Ghojogh, B., Ghodsi, A., Karray, F., & Crowley, M. (2021). Generative adversarial networks and adversarial autoencoders: Tutorial and survey. arXiv preprint arXiv:2111.13282. https://doi.org/10.1007/978-3-031-10602-6_21

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems, 27*.

Gray, J. (2023). Copyright eligibility for AI-generated images: the threshold of human creativity. *Gonz. L. Rev., 59*, 507.

Gui, J., Sun, Z., Wen, Y., Tao, D., & Ye, J. (2021). A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE transactions on knowledge and data engineering, 35*(4), 3313-3332. https://doi.org/10.1109/TKDE.2021.3130191

Haase, J., & Hanel, P. H. P. (2023). Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity. arXiv. arXiv preprint arXiv:2303.12003. https://doi.org/10.1016/j.yjoc.2023.100066

Hamilton, M. (2020). *The coded gaze: Unpacking biases in algorithms that perpetuate inequity.* The Rockefeller Foundation.

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation, 16*(1), 45-74. https://doi.org/10.1007/s12559-023-10179-8

He, R., Cao, J., & Tan, T. (2025). Generative artificial intelligence: a historical perspective. *National Science Review, 12*(5), nwaf050. https://doi.org/10.1093/nsr/nwaf050

Hemraj, S. (2025). AI and the future of creative development: redefining digital media production. *AI and Ethics,* 1-15. https://doi.org/10.1007/s43681-025-00765-x

Hendrickson, L. (2025, February 5). Protecting Public Figures and Artists' Likeness in the Age of AI. Identity. https://www.identity.com/protecting-public-figures-and-artists-likeness-in-the-age-of-ai/

Horton, A. (2024, July 27). *"Hold on to your seats": how much will AI affect the art of film-making?*. The Guardian. https://www.theguardian.com/film/article/2024/jul/27/artificial-intelligence-movies

Huang, D., Yan, C., Li, Q., & Peng, X. (2024). From large language models to large multimodal models: A literature review. *Applied Sciences, 14*(12), 5068. https://doi.org/10.3390/app14125068

Huang, J., Hitchen, G., & Dogan, S. (2025). Generative AI in the screen and live performance industries: A conceptual framework and prospects for future research. *Convergence,* 13548565251383409. https://doi.org/10.1177/13548565251383409

Huang, Y., Lv, S., Tseng, K. K., Tseng, P. J., Xie, X., & Lin, R. F. Y. (2023). Recent advances in artificial intelligence for video production system. *Enterprise Information Systems, 17*(11), 2246188. https://doi.org/10.1080/17517575.2023.2246188

Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks–enabled human–artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. *Frontiers in artificial intelligence, 4*, 604234. https://doi.org/10.3389/frai.2021.604234

Hutson, J. (2024). Art in the Age of Virtual Reproduction. In *Art and Culture in the Multiverse of Metaverses: Immersion, Presence, and Interactivity in the Digital Age* (pp. 55-98). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-66320-8_3

Hutson, J., & Smith, A. (2025). Cinematic Algorithms: The Rise of Generative AI in Video Art and Visual Culture. https://doi.org/10.1201/9781003601128

Jabbar, A., Li, X., & Omar, B. (2021). A survey on generative adversarial networks: Variants, applications, and training. *ACM Computing Surveys (CSUR), 54*(8), 1-49. https://doi.org/10.1145/3463475

Jin, Y., Wang, Y., Long, M., Wang, J., Yu, P. S., & Sun, J. (2020, July). A multi-player minimax game for generative adversarial networks. In *2020 IEEE International Conference on Multimedia and Expo* (ICME) (pp. 1-6). IEEE. https://doi.org/10.1109/ICME46284.2020.9102779

Joshi, S. (2025). Introduction to Diffusion Models, Autoencoders and Transformers: *Review of Current Advancements.* https://doi.org/10.20944/preprints202503.1431.v1

Kang, J., Poria, S., & Herremans, D. (2024). Video2music: Suitable music generation from videos using an affective multimodal transformer model. *Expert Systems with Applications, 249,* 123640. https://doi.org/10.1016/j.eswa.2024.123640

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., & Aila, T. (2020). Training generative adversarial networks with limited data. *Advances in neural information processing systems, 33*, 12104-12114.

Kingma, D. P., & Welling, M. (2013, December). Auto-encoding variational bayes.

Kumar, S., Datta, S., Singh, V., Datta, D., Singh, S. K., & Sharma, R. (2024). Applications, challenges, and future directions of human-in-the-loop learning. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3401547

Laco, V. E. (2024). "As If It Was A Funhouse Mirror": Investigating Consumers' Attitudes Towards AI-Generated Videos.

Lee, E. (2024). Prompting progress: authorship in the age of AI. *Fla. L. Rev., 76,* 1445. https://doi.org/10.2139/ssrn.4609687

Leininger, P., Weber, C. J., & Rothe, S. (2025, June). Understanding Creative Potential and Use Cases of AI-Generated Environments for Virtual Film Productions: Insights from Industry Professionals. In *Proceedings of the 2025 ACM International Conference on Interactive Media Experiences* (pp. 60-78). https://doi.org/10.1145/3706370.3727853

Liang, C. X., Tian, P., Yin, C. H., Yua, Y., An-Hou, W., Ming, L., ... & Liu, M. (2024). A Comprehensive Survey and Guide to Multimodal Large Language Models in Vision-Language Tasks. arXiv preprint arXiv:2411.06284.

Lin, C. Y., Tseng, T. L., Emon, S. H., & Tsai, T. H. (2025). Generative AI-Driven Data Augmentation for Robust Virtual Metrology: GANs, VAEs, and Diffusion Models. IEEE Transactions on Semiconductor Manufacturing. https://doi.org/10.1109/TSM.2025.3569229

Liu, Q., Zhu, J., Yang, Y., Dai, Q., Du, Z., Wu, X. M., ... & Dong, Z. (2024, August). Multimodal pretraining, adaptation, and generation for recommendation: A survey. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 6566-6576). https://doi.org/10.1145/3637528.3671473

Luk, C., Chung, H. L., Yim, W. K., & Leung, C. W. (2024). Regulating generative AI: Ethical considerations and explainability benchmarks. https://doi.org/10.31219/osf.io/h74gw

Mainini, J. (2023). Stephen Thaler v. Shira Perlmutter, et al.-No. 22-1564 (BAH)(DDC 2023). Intell. Prop. & Tech. LJ, 28, 65.

Mathewson, K. W. (2019). Humour-in-the-loop: Improvised Theatre with Interactive Machine Learning Systems.

Miller, J. (2024). *Do androids dream of electric screens?* The effects of AI on the film industry.

Mirowski, P., Mathewson, K. W., Pittman, J., & Evans, R. (2023, April). Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals. In *Proceedings of the 2023 CHI conference on human factors in computing systems* (pp. 1-34). https://doi.org/10.1145/3544548.3581225

Mitra, R., & Zualkernan, I. (2025). Music Generation Using Deep Learning and Generative AI: A Systematic Review. IEEE Access. https://doi.org/10.1109/ACCESS.2025.3531798

Mittal, S., Joshi, M., Vats, P., Upadhayay, G. M., Vats, S. K., & Kumar, S. (2024, March). Virtual Illusions: Unleashing Deepfake Expertise for Enhanced Visual Effects in Film Production. In *2024 11th International Conference on Reliability, Infocom Technologies and Optimization* (Trends and Future Directions)(ICRITO) (pp. 1-6). IEEE. https://doi.org/10.1109/ICRITO61523.2024.10522334

Mozelius, P., & Humble, N. (2024, July). On the use of generative AI for literature reviews: an exploration of tools and techniques. In *Proceedings of the 23rd European Conference on Research Methodology for Business and Management Studies* (ECRM 2024). Mid Sweden University, University of Gävle, & Uppsala University. https://doi.org/10.34190/ecrm.23.1.2528

Ni, M., Fan, Y., Zhang, L., & Zuo, W. (2024). Visual-o1: Understanding ambiguous instructions via multi-modal multi-turn chain-of-thoughts reasoning. arXiv preprint arXiv:2410.03321.

Ni, M., Wu, C., Yuan, H., Yang, Z., Gong, M., Wang, L., ... & Duan, N. (2024). AutoDirector: Online Auto-scheduling Agents for Multi-sensory Composition. arXiv preprint arXiv:2408.11564.

Oberting IV, V. A. (2024). Generative Artificial Intelligence and Copyright in the Film and Media Industry. *Washington and Lee Law Review Online, 82*(2), 123.

Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. (2012). Qualitative analysis techniques for the review of the literature. *Qualitative report, 17*, 56.

Oshiesh, J. A. R. (2025). The Poetics of Code: Generative AI and the Redefinition of Literary Creativity. *The Voice of Creative Research, 7*(1), 195-212. https://doi.org/10.53032/tvcr/2025.v7n1.23

Park, H., Jung, M. W. R., Rakhmonov, S., & Kim, S. (2024, December). Ourmuse: Plot-Specific AI Music Generation for Video Advertising. In *2024 IEEE International Conference on Big Data* (BigData) (pp. 3200-3203). IEEE. https://doi.org/10.1109/BigData62323.2024.10826083

Pataranutaporn, P., Danry, V., Leong, J., Punpongsanon, P., Novy, D., Maes, P., & Sra, M. (2021). AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence, 3*(12), 1013-1022. https://doi.org/10.1038/s42256-021-00417-9

Phillips, D., Brice, D., Phillips, E., Lydon, D., & Chang, J. R. (2024). AI screen skills study.

Przegalinska, A., & Triantoro, T. (2024). *Converging minds: The creative potential of collaborative AI.* CRC Press. https://doi.org/10.1201/9781032656618

Ramalakshmi, S., & Asha, G. (2024). Exploring generative AI: Models, applications, and challenges in data synthesis. *Asian Journal of Research in Computer Science, 17*(12), 123-136. https://doi.org/10.9734/ajrcos/2024/v17i12533

Ren, Z., Jin, M., Nie, H., Shen, J., Dong, A., & Zhang, Q. (2025). Towards Realistic Human Motion Prediction with Latent Diffusion and Physics-Based Models. *Electronics, 14*(3), 605. https://doi.org/10.3390/electronics14030605

Rezek, A. (2024). Filling the Enforcement Gap: Alternative Dispute Resolution as an Approach to Solving" Copyright" Disputes For AI-Generated Content. *J. Disp. Resol., 190.*

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695). https://doi.org/10.1109/CVPR52688.2022.01042

Rosati, E. (2025). The future of the movie industry in the wake of generative AI: A perspective under EU and UK copyright law. *Computer Law & Security Review, 59*, 106207. https://doi.org/10.1016/j.clsr.2025.106207

Rozas, L. W., & Klein, W. C. (2010). The value and purpose of the traditional qualitative literature review. *Journal of evidence-based social work, 7*(5), 387-399. https://doi.org/10.1080/15433710903344116

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-based systems, 263*, 110273. https://doi.org/10.1016/j.knosys.2023.110273

Santoso, B., & Wijayanti, R. (2024). Human-AI Collaboration in Creative Industries: Workflows in Media Production and Community-Driven Platforms. *Transactions on Artificial Intelligence, Machine Learning, and Cognitive Systems, 9*(11), 11-26.

Schryen, G. (2015). Writing qualitative is literature reviews—guidelines for synthesis, interpretation, and guidance of research. *Communications of the Association for Information Systems, 37*(1), 12. https://doi.org/10.17705/1CAIS.03712

Serena, A. (2025). *The Impact of Generative AI on the Film Industry: Opportunities, Challenges, and Ethical Implications* (Doctoral dissertation, Politecnico di Torino).

Seshadri, N. (2021). The Dilemma of Artificial Intelligence Generated Works and Indian Copyright Law. *Indian JL & Legal Rsch., 3*, 1.

Shao, H., Yao, S., Sun, D., Zhang, A., Liu, S., Liu, D., ... & Abdelzaher, T. (2020, November). Controlvae: Controllable variational autoencoder. In *International conference on machine learning* (pp. 8655-8664). PMLR.

Sharma, P., Kumar, M., Sharma, H. K., & Biju, S. M. (2024). Generative adversarial networks (GANs): introduction, taxonomy, variants, limitations, and applications. *Multimedia Tools and Applications,* 1-48. https://doi.org/10.1007/s11042-024-18767-y

Sinha, S., Datta, S., Kumar, R., Bhattacharya, S., Sarkar, A., & Das, K. (2024). Exploring Creativity: The Development and Uses of Generative AI. In *The Pioneering Applications of Generative AI* (pp. 167-198). IGI Global. https://doi.org/10.4018/979-8-3693-3278-8.ch008

Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. (2015, June). Deep unsupervised learning using nonequilibrium thermodynamics. In International conference on machine learning (pp. 2256-2265). pmlr.

Solanki, S. R., & Khublani, D. K. (2024). From Script to Screen: Unveiling Text-to-Video Generation. In *Generative Artificial Intelligence: Exploring the Power and Potential of Generative AI* (pp. 81-112). Berkeley, CA: Apress. https://doi.org/10.1007/979-8-8688-0403-8_3

Somaini, A. (2023). Algorithmic images: Artificial intelligence and visual culture. *Grey Room,* (93), 74-115. https://doi.org/10.1162/grey_a_00383

Song, T., Sun, J., Chen, B., Peng, W., & Song, J. (2019). Latent space expanded variational autoencoder for sentence generation. IEEE Access, 7, 144618-144627. https://doi.org/10.1109/ACCESS.2019.2944630

Swarnakar, S. (2024). A*rtificial Intelligence and Cinema-Exploring the Implications of Artificial Intelligence in Cinema.* The Media Mosaic: Exploring Diverse Artistic Forms, 21-26.

Thalpage, N. (2023). Unlocking the black box: Explainable artificial intelligence (XAI) for trust and transparency in ai systems. *J. Digit. Art Humanit, 4*(1), 31-36. https://doi.org/10.33847/2712-8148.4.1_4

The Frost - Waymark. (2024). Thefrost.ai. https://www.thefrost.ai/

Totlani, K. (2023). The evolution of generative AI: Implications for the media and film industry. *International Journal for Research in Applied Science and Engineering Technology, 11*(10), 973-980. https://doi.org/10.22214/ijraset.2023.56140

Townsend, S. (2024, March 2). Exploring the Impact of AI on Film Production in 2024. Medium. https://medium.com/@channelasaservice/exploring-the-impact-of-ai-on-film-production-in-2024-f02da745af00

Trees, J. (2024). Artistry in Flux: Navigating the Nuances of Generative AI's Impact on Creativity in Art and Film.

Truong, V. T., Dang, L. B., & Le, L. B. (2025). Attacks and defenses for generative diffusion models: A comprehensive survey. *ACM Computing Surveys, 57*(8), 1-44. https://doi.org/10.1145/3721479

Uddin, S. M. I., Sumon, R. I., Mozumder, M. A. I., Chowdhury, M. K. H., Armand, T. P. T., & Kim, H. C. (2025). Innovations and Challenges of AI in Film: A Methodological Framework for Future Exploration. *ACM Transactions on Multimedia Computing, Communications and Applications.* https://doi.org/10.1145/3736724

Varghese, M. J. (2023). T*he AI Copyright Conundrum: Navigating Authorship and Ownership in the AI Landscape.* Jus Corpus LJ, 4, 309.

Verma, K. (2024). Digital Deception: The Impact of Deepfakes on Privacy Rights. *Lex Scientia Law Review, 8*(2), 859-896. https://doi.org/10.15294/lslr.v8i2.13749

Viktoriia, M. (2024). Generative AI for 2.5 D Content Creation with Depth-Guided Object Placement.

Von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology, 34*(4), 1607-1622. https://doi.org/10.1007/s13347-021-00477-0

Wang, K., Deng, S., Shi, J., Hatzinakos, D., & Tian, Y. (2024). AV-DiT: Efficient Audio-Visual Diffusion Transformer for Joint Audio and Video Generation. arXiv preprint arXiv:2406.07686. https://doi.org/10.1145/3746027.3755713

Wang, Y. (2023). Artificial Creativity-Ethical Reflections on AI's Role in Artistic Endeavors. Authorea Preprints. https://doi.org/10.36227/techrxiv.23897169.v1

Washington, J. (2023). The impact of generative artificial intelligence on writer's self-efficacy: a critical literature review. Available at SSRN 4538043. https://doi.org/10.2139/ssrn.4538043

Xing, Z., Feng, Q., Chen, H., Dai, Q., Hu, H., Xu, H., ... & Jiang, Y. G. (2024). A survey on video diffusion models. *ACM Computing Surveys, 57*(2), 1-42. https://doi.org/10.1145/3696415

Xu, Y. (2025). Balancing creativity and automation: The influence of AI on modern film production and dissemination. arXiv preprint arXiv:2504.19275. https://arxiv.org/abs/2504.19275

Yang, S., Ge, Y., Li, Y., Chen, Y., Ge, Y., Shan, Y., & Chen, Y. (2024). Seed-story: Multimodal long story generation with large language model. arXiv preprint arXiv:2407.08683.

Yin, M. Machine Learning in Storyboarding and Storytelling (Doctoral dissertation, Toronto Metropolitan University).

Yu, L., Sun, X., Zhou, W., & Gabbouj, M. (2025). Text-Audio-Visual-conditioned Diffusion Model for Video Saliency Prediction. arXiv preprint arXiv:2504.14267.

Zhang, H., Li, X., & Bing, L. (2023). Video-llama: An instruction-tuned audio-visual language model for video understanding. arXiv preprint arXiv:2306.02858. https://doi.org/10.18653/v1/2023.emnlp-demo.49

Zhong, C., Yu, J., Cao, Y., Wu, S., Wu, W., & Zhang, K. (2024). SoundScape: A Human-AI Co-Creation System Making Your Memories Heard. arXiv preprint arXiv:2410.08136.

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS nexus, 3*(3), 052. https://doi.org/10.1093/pnasnexus/pgae052