IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

THE ROLE OF ARTIFICIAL INTELLIGENCE IN ENHANCING THE QUALITY OF CGI AND VISUAL EFFECTS IN MOVIES

Author:[1] Viknesh Raja R S*,[2] Dr B Senthil Kumar

[1] Research Scholar, [2]Assistant Professor

[1]VELS Institute of Science, Technology and Advanced Studies(VITAS), Chennai

Abstract: In films, Visual Effects (VFX) are meticulously crafted by skilled artists who painstakingly create and manipulate computer-generated imagery (CGI) to enhance a film's storytelling and visual aesthetics. However, with the rise of artificial intelligence (AI), the landscape of VFX is transforming. Consequently, the widespread implementation of AI has sparked concerns within the film industry about the potential replacement of human contributions, akin to trends observed in other sectors previously dominated by humans. By enhancing the creation of computer-generated imagery (CGI) and visual effects (VFX), artificial intelligence (AI) is revolutionizing the film business. In order to produce realistic individuals, settings, animals, action scenes, and fantastical worlds, CGI and VFX are crucial technology. In the past, these procedures needed costly software, highly qualified artists, and a significant amount of manufacturing time. New possibilities for enhancing these operations have been made possible by the advancements in machine learning, deep learning, generative adversarial networks (GANs), diffusion models, and neural rendering. This essay looks at how AI can improve CGI and visual effects in motion pictures. Character development, facial animation, motion capture, environment construction, picture enhancement, compositing, and post-production are among the applications covered. The benefits, drawbacks, moral dilemmas, and potential applications of AI-assisted filmmaking are also discussed in the article.

Index Terms **-** Artificial Intelligence, CGI, Visual Effects, Machine Learning, Deep Learning, Generative AI, Film Industry, Animation, Post-production.

I. INTRODUCTION

The film industry has undergone remarkable technological transformation over the last few decades, particularly through the development of Computer-Generated Imagery (CGI) and advanced Visual Effects (VFX). These technologies have revolutionized the way filmmakers create, manipulate, and present visual narratives. CGI enables artists to design realistic or imaginative characters, objects, environments, and special effects that would be difficult, costly, or impossible to achieve through conventional filmmaking techniques. As a result, CGI has become an essential component of contemporary cinema, expanding the creative possibilities available to filmmakers and enhancing audience engagement.

Visual effects play a crucial role in modern filmmaking, especially in genres such as science fiction, fantasy, superhero, animation, historical, and action films. Through the integration of digital technologies, filmmakers can construct immersive cinematic worlds and create visually compelling experiences that transcend the limitations of physical production. Prince (2012) argues that digital visual effects have become a significant means of constructing cinematic reality, enabling filmmakers to represent events, spaces, and characters that extend beyond the capabilities of traditional cinematography. Consequently, CGI and VFX have become indispensable tools for storytelling in the digital age.

In recent years, Artificial Intelligence (AI) has emerged as another transformative force within the film industry. AI technologies possess the ability to analyse large volumes of visual data, recognise patterns, generate content, and automate complex workflows. According to Tsiavos and Kitsios (2025), AI is increasingly influencing every stage of the film industry's value chain, including creation, production, distribution, and exhibition, while simultaneously raising important questions concerning authorship, creative integrity, and labour displacement. The growing adoption of generative AI technologies has further accelerated discussions regarding the relationship between human creativity and machine-assisted production.

The integration of AI into CGI and VFX production has introduced new opportunities for efficiency and innovation. Traditionally, visual-effects artists spent considerable time performing labour-intensive tasks such as rot scoping, object tracking, motion capture refinement, image segmentation, facial animation, and background generation. AI-powered tools can now automate many of these processes, significantly reducing production time while maintaining high levels of visual quality. Recent studies indicate that AI-assisted workflows enhance productivity by supporting artists in technical tasks, allowing them to focus more on creative decision-making and artistic expression.

Despite these advantages, scholars caution against viewing AI as a complete replacement for human creativity. Research on contemporary VFX practices suggests that human expertise, artistic judgment, and creative vision remain central to the filmmaking process. AI functions most effectively as a collaborative tool that augments human capabilities rather than replacing them entirely. Narayan et al. (2026) argue that while AI may automate certain technical aspects of visual-effects production, the demand for skilled creative labour continues to persist. Similarly, discussions on AI-generated media emphasize that meaningful storytelling, emotional depth, and cultural understanding still rely heavily on human creators.

Therefore, AI should be understood not merely as a substitute for traditional CGI and VFX techniques but as a creative and technical assistant that enhances artistic workflows. By combining human imagination with computational intelligence, filmmakers can achieve increasingly sophisticated visual results while preserving the creative essence of cinematic storytelling. As AI technologies continue to evolve, the future of CGI and VFX is likely to be characterized by collaborative human–AI partnerships that balance technological innovation with artistic authenticity.

2. Evolution of CGI and Visual Effects

Visual effects (VFX) have been an integral part of filmmaking since the early development of cinema. Traditional techniques such as miniatures, matte paintings, stop-motion animation, optical effects, and practical effects enabled filmmakers to create imaginative scenes beyond the limitations of live-action production. The emergence of computer-generated imagery (CGI) significantly transformed the industry by allowing digital elements to be seamlessly integrated with live-action footage, leading to increasingly sophisticated visual storytelling. As Dixon (2019) observes, CGI has become a defining feature of modern "synthetic cinema," reshaping contemporary filmmaking practices.

Artificial Intelligence (AI) represents the latest advancement in the evolution of visual effects. Unlike conventional CGI, which requires extensive manual input from artists, AI systems can learn patterns from large datasets and automate various creative and technical processes. Recent studies indicate that AI is increasingly influencing all stages of film production, including visual-effects creation, animation, post-

production, and content generation (Tsiavos & Kitsios, 2025). Modern deep-learning and generative AI models can analyse images and videos, automate visual tasks, and generate realistic digital content, thereby expanding the creative possibilities of CGI and VFX. Consequently, AI is emerging as a valuable tool that enhances efficiency while supporting innovation in contemporary filmmaking.

3. Role of AI in CGI and Visual Effects

3.1 AI-Based Image Processing

Image processing has become an important application of Artificial Intelligence (AI) in contemporary visual-effects production. AI-based systems can analyse digital images, identify objects and faces, detect patterns, segment backgrounds, and extract visual features with increasing accuracy. These capabilities support filmmakers and visual-effects artists in processing and enhancing large volumes of digital footage. Archana and Jeevaraj (2024) explain that deep-learning techniques have significantly advanced image-processing tasks such as denoising, enhancement, segmentation, feature extraction, and classification.Recent developments in computer vision have further strengthened the role of AI in visual processing. Zhao et al. (2024) note that deep-learning models, particularly convolutional neural networks, have improved applications such as object detection, image classification, segmentation, and image super-resolution. Similarly, Trigka and Dritsas (2025) highlight the growing ability of deep-learning systems to process complex visual information and improve the efficiency and accuracy of image-processing operations.

AI-based image processing is particularly useful in film and television production, where visual quality is essential for creating convincing cinematic environments. Liu (2025) demonstrates the application of AI models to film-image processing, including image denoising and detail restoration. Such techniques can help improve footage before it is integrated with CGI and VFX elements. AI can also assist in restoring damaged images, enhancing resolution, reducing visual noise, and improving the consistency of digitally generated scenes.The continuing development of deep-learning architectures is expanding these possibilities. Xiong, Tang, and Verbeek (2026) observe that modern deep-learning systems increasingly combine automated feature extraction with task-specific image analysis, reducing the need for extensive manual processing. These developments are particularly relevant to VFX workflows, where large quantities of visual data must be analysed and modified efficiently.

Therefore, AI-based image processing has become an important technological component of modern CGI and VFX production. It can reduce repetitive technical work, improve image quality, and support faster visual development while allowing artists to concentrate on creative and aesthetic decisions. The integration of AI with human expertise consequently provides new opportunities for efficient and sophisticated visual storytelling.

3.2 AI and Rotoscoping

Rotoscoping is an important technique in visual effects (VFX) used to separate a foreground subject, such as a person or object, from its background. Traditionally, this process involves creating and refining masks across multiple frames, making it a time-consuming and repetitive task for VFX artists (Adobe, 2026). The development of AI-based image segmentation has introduced more efficient approaches to subject isolation by automatically identifying objects and generating segmentation masks (Kirillov et al., 2023). Advanced models such as SAM 2 extend these capabilities to video by maintaining segmentation across frames and reducing the need for repeated user interaction (Ravi et al., 2024). Research on video segmentation further demonstrates the growing application of foundation models in video editing and visual-processing tasks (Zhang et al., 2024). Similarly, AI-powered rotoscoping tools such as Adobe's Roto Brush use machine learning to track subjects and propagate selections across frames, thereby

reducing extensive frame-by-frame manual work (Adobe, 2020). AI therefore supports VFX artists by automating repetitive segmentation and tracking tasks while allowing them to focus on refining edges, correcting errors, and achieving accurate compositing. These capabilities are particularly useful in large-scale productions where numerous VFX shots require efficient and consistent processing.

3.3 AI and Facial Animation

Facial animation is one of the most complex stages of digital character production because realistic digital performers must reproduce subtle facial movements, eye movements, lip synchronization, and emotional expressions. Traditional facial animation relies on techniques such as keyframe animation, facial rigging, motion capture, and extensive manual refinement. The development of artificial intelligence (AI) and machine-learning techniques has introduced new methods for analysing facial movement and transferring human performance to digital characters. Thies et al. (2016), for example, demonstrated real-time facial capture and reenactment by tracking facial expressions and transferring them between video sequences.

Advances in three-dimensional face reconstruction have further improved the ability of computational systems to estimate facial identity, geometry, pose, and expression from visual information. Feng et al. (2021) demonstrated the learning of detailed and animatable 3D face models from images captured in uncontrolled environments. Similarly, neural-rendering research has combined computer-graphics techniques with learned representations to produce realistic digital faces and performances (Tewari et al., 2020). These developments have expanded performance-driven facial animation by enabling systems to learn relationships between human facial expressions and their digital representations.

Transformer-based approaches have further extended AI-assisted facial animation. Fan et al. (2022), through FaceFormer, demonstrated speech-driven 3D facial animation using a transformer-based autoregressive model capable of generating sequences of animated facial meshes from audio input. Such systems have the potential to reduce repetitive manual operations and provide animators with an initial performance that can subsequently be adjusted. AI-assisted facial animation can therefore be regarded as a collaborative process in which computational systems support facial analysis and animation while human artists retain responsibility for expression, timing, emotional interpretation, and visual continuity.

4. AI IN CHARACTER CREATION AND ANIMATION

Artificial intelligence can contribute to several stages of digital character production, including conceptualisation, modelling, texturing, rigging, animation, and rendering. The emergence of generative models has expanded the capacity of computational systems to produce visual content from textual, visual, or multimodal instructions. Generative adversarial networks (GANs), introduced by Goodfellow et al. (2014), established an influential framework for synthesising visual content through the interaction of generative and discriminative networks. Subsequent developments in GAN architectures significantly improved the quality and realism of generated images (Karras et al., 2019; Karras et al., 2020).

Transformer architectures have also influenced the development of generative and multimodal AI systems. Vaswani et al. (2017) introduced the transformer architecture, which subsequently became an important foundation for modern generative systems. Radford et al. (2021) demonstrated the effectiveness of large-scale visual-language learning, while diffusion-based approaches introduced new methods for high-quality image synthesis and manipulation (Ho et al., 2020; Dhariwal & Nichol, 2021; Rombach et al., 2022). Comprehensive research on diffusion models indicates their expanding applications in image generation, editing, video synthesis, and related visual tasks (Yang et al., 2024).

Within character production, these technologies can assist artists in generating concept sketches, costumes, facial references, textures, and alternative character designs. However, generated content normally requires human assessment and refinement before being incorporated into a professional production pipeline. AI therefore functions more appropriately as a creative and technical support system than as a complete substitute for modelling, animation, and artistic expertise.

Human-pose estimation represents another significant application of AI in character animation. Deep-learning research has produced substantial improvements in estimating two-dimensional and three-dimensional human poses from images and video (Dang et al., 2019; Chen et al., 2020; Zheng et al., 2021). Open Pose demonstrated the practical application of neural networks for detecting human body key points in images and video (Cao et al., 2017). Later approaches improved the accuracy and robustness of pose estimation through high-resolution representations and improved feature learning (Xiao et al., 2018; Sun et al., 2019). These developments provide important technological foundations for AI-assisted character animation and motion capture.

5. AI FOR ENVIRONMENT AND BACKGROUND CREATION

Digital environments are essential components of contemporary CGI and VFX production. Films may require realistic urban spaces, historical settings, imaginary worlds, extra-terrestrial landscapes, underwater locations, or futuristic environments. Traditionally, the construction of such environments involves modelling, texturing, lighting, simulation, compositing, and rendering. AI-based techniques can support these processes by generating preliminary visual concepts and assisting artists in exploring alternative environments.

Generative models have demonstrated considerable potential for image synthesis and visual transformation (Brock et al., 2019; Ho et al., 2020; Rombach et al., 2022). Text-to-image systems have further demonstrated the ability to transform natural-language descriptions into visual representations, thereby supporting rapid concept development (Ramesh et al., 2021; Saharia et al., 2022). Such systems can be particularly useful during pre-production because filmmakers can explore multiple environmental possibilities before investing substantial resources in detailed three-dimensional construction.

Neural rendering provides another important technological foundation for AI-assisted environment creation. Tewari et al. (2020) describe neural rendering as an emerging area that combines computer graphics with machine-learning techniques for the synthesis and manipulation of visual content. Mildenhall et al. (2020) introduced Neural Radiance Fields (NeRF), demonstrating how neural representations can reconstruct and render three-dimensional scenes from multiple images. Subsequently, Kerbl et al. (2023) introduced 3D Gaussian Splatting as an efficient approach for real-time radiance-field rendering.

Despite these developments, AI-generated environments require careful human supervision. Generated content may contain spatial inconsistencies, inaccurate architectural features, inappropriate lighting, or elements that conflict with the narrative setting. VFX artists and filmmakers must therefore evaluate AI-generated material for continuity, spatial coherence, historical accuracy, lighting consistency, and artistic suitability.

6. AI IN MOTION CAPTURE AND VISUAL EFFECTS

Motion capture enables filmmakers to record human performances and transfer them to digital characters. AI can enhance this process through automated body-pose estimation, joint detection, movement tracking, and motion reconstruction from video. Deep-learning-based human-pose estimation has consequently

become an important area of computer-vision research with applications in animation, augmented reality, virtual reality, and human-computer interaction (Dang et al., 2019; Zheng et al., 2021).

Open Pose demonstrated the ability to estimate human body and facial key points from visual input (Cao et al., 2017). DensePose extended this approach by establishing dense correspondence between human images and a three-dimensional body surface, allowing more detailed representations of human movement and appearance (Güler et al., 2018). Similarly, Pavllo et al. (2019) demonstrated that temporal information from video can contribute to accurate three-dimensional human-pose estimation.

AI-assisted performance tracking can also facilitate the transfer of human performances to digital characters. Face2Face demonstrated real-time facial capture and reenactment from RGB video, showing how facial expressions can be tracked and transferred between performers and target faces (Thies et al., 2016). Deep Video Portraits further demonstrated the use of deep-learning methods for transferring human appearance and facial performance (Kim et al., 2018). These developments can reduce some manual processing requirements and allow animators to focus more extensively on performance refinement and artistic interpretation.

More recent foundation models have expanded automated visual tracking and segmentation. Segment Anything Model 2 (SAM 2), developed by Ravi et al. (2024), provides promptable segmentation for images and videos and incorporates memory mechanisms for tracking objects across video sequences. Such capabilities have potential applications in VFX workflows involving subject isolation, masking, object tracking, and compositing.

7. AI IN POST-PRODUCTION

Post-production involves editing, compositing, colour correction, image enhancement, visual-effects integration, and other processes required to produce the final audiovisual work. AI can support these activities by analysing visual information, identifying objects and faces, tracking movement, segmenting images, and automating selected repetitive operations.

Deep-learning techniques have contributed substantially to developments in image recognition, object detection, and semantic segmentation. Residual neural networks improved the extraction of hierarchical visual features from images (He et al., 2016), while Faster R-CNN provided an effective framework for object detection and localisation (Ren et al., 2015). Fully convolutional networks further established important methods for semantic image segmentation (Long et al., 2015).

Image segmentation is particularly relevant to VFX because it enables foreground objects, performers, and backgrounds to be separated for further manipulation. Kirillov et al. (2023) introduced the Segment Anything Model (SAM), demonstrating the potential of general-purpose segmentation models. Ravi et al. (2024) extended this direction through SAM 2, which supports promptable segmentation in both images and videos. These technologies can assist with masking, rotoscoping, object tracking, and compositing tasks.

Recent research into AI-assisted filmmaking indicates that deep-learning systems are increasingly being investigated across several stages of cinematic production, including image processing, video generation, editing, and cinematographic decision-making (Azzarelli et al., 2025). AI can therefore function as an augmentative technology within post-production, assisting professionals while leaving final creative decisions under human control.

8. ADVANTAGES OF AI IN CGI AND VFX

8.1 Reduction in Production Time

One of the significant potential benefits of AI-assisted VFX is the reduction of time spent on repetitive technical processes. Automated segmentation, object tracking, pose estimation, and image processing can reduce the number of manual operations required from artists (Kirillov et al., 2023; Ravi et al., 2024). This can allow production teams to allocate more time to creative refinement and quality control.

8.2 Improved Visual Processing

AI systems can analyse large quantities of visual information and identify patterns, objects, and relationships within images and video. Developments in neural rendering, image synthesis, and computer vision have introduced new methods for processing and manipulating digital imagery (Tewari et al., 2020; Yang et al., 2024). These technologies can support visual enhancement and the creation of complex digital scenes.

8.3 Reduction of Repetitive Work

Many VFX processes involve repetitive activities such as masking, tracking, segmentation, and preliminary animation. AI can assist with selected portions of these tasks, thereby reducing repetitive workloads. The resulting workflow enables artists to devote greater attention to creative interpretation, visual consistency, and quality assurance.

8.4 Greater Creative Possibilities

Generative AI provides filmmakers and artists with the ability to explore numerous visual possibilities within a relatively short period. GANs, transformers, and diffusion models have expanded the range of computational techniques available for visual generation and manipulation (Goodfellow et al., 2014; Vaswani et al., 2017; Ho et al., 2020; Rombach et al., 2022). Such technologies can support experimentation during concept development and pre-production.

8.5 Increased Accessibility

AI-assisted tools may reduce the technical complexity involved in selected visual-production tasks. This can provide smaller production teams and independent creators with opportunities to experiment with techniques that traditionally required specialised expertise and extensive resources. However, the quality of the final output continues to depend on appropriate artistic judgement and technical supervision.

9. CHALLENGES AND ETHICAL ISSUES

9.1 Employment and Creative Labour

The increasing automation of creative and technical processes has generated discussion about the changing roles of artists, technicians, performers, editors, and other media professionals. Research on AI and cinematographic production identifies both opportunities for automation and concerns regarding the transformation of creative labour (Azzarelli et al., 2025). The central issue is therefore not simply whether AI can perform particular tasks, but how its adoption may alter professional responsibilities, required skills, employment structures, and remuneration.

AI-assisted production may consequently require new forms of professional expertise, including AI supervision, prompt development, data management, visual evaluation, and quality control. The future production environment may therefore involve a combination of computational automation and specialised human expertise.

9.2 Copyright and Authorship

Generative AI has introduced complex questions concerning copyright, authorship, training data, ownership, and the legal status of AI-generated outputs. Oberting (2024) discusses the implications of generative AI for copyright and authorship within the film and media industry. Similarly, Mora and Llamas (2025) examine the relationship between generative AI and the movie industry under European Union and United Kingdom copyright frameworks, particularly in relation to training data, originality, authorship, and liability.

These concerns are particularly significant in film production because audiovisual works contain numerous copyrighted elements, including images, performances, recordings, scripts, and visual assets. Consequently, responsible AI-assisted production requires careful consideration of licensing, attribution, documentation, authorship, and human creative contribution.

9.3 Human Creativity

Generative AI can produce visually sophisticated material, but technical quality does not necessarily ensure narrative coherence, emotional meaning, cultural sensitivity, or artistic originality. Research on foundation models has highlighted both their broad capabilities and their limitations concerning data, evaluation, bias, and controllability (Bommasani et al., 2021). Similarly, Bender et al. (2021) identify important concerns concerning the limitations and social implications of large-scale AI systems.

Human creativity therefore remains essential in determining whether generated content is appropriate for a particular narrative, character, cultural context, audience, and visual style. AI can provide alternatives and accelerate experimentation, but creative interpretation and final artistic judgement remain important components of filmmaking.

9.4 Digital Humans and Performer Rights

The development of AI-generated digital humans raises questions concerning identity, consent, likeness, and performer rights. Technologies such as facial reenactment have demonstrated the ability to manipulate and reproduce human facial performances (Thies et al., 2016; Kim et al., 2018). As digital doubles, synthetic performances, and other forms of virtual representation become more sophisticated, issues concerning consent and control over an individual's likeness become increasingly significant.

Responsible use of digital-human technologies therefore requires appropriate consideration of consent, contractual arrangements, attribution, identity protection, and the intended use of synthetic representations. These issues demonstrate that the development of AI in CGI and VFX involves not only technological considerations but also legal, professional, and ethical responsibilities.

10. FUTURE SCOPE

The future development of AI-assisted CGI and VFX is likely to involve greater integration between generative models, computer vision, neural rendering, multimodal systems, and conventional graphics pipelines. Recent advances in diffusion models, neural rendering, segmentation, and video processing indicate a rapidly developing technological landscape (Tewari et al., 2020; Yang et al., 2024; Ravi et al., 2024).

Future production systems may allow filmmakers to describe preliminary characters, environments, camera movements, or visual effects through natural-language instructions and subsequently refine the generated material using conventional production software. Multimodal foundation models may further support workflows that integrate text, images, audio, video, and other forms of information (Bommasani et al., 2021).

Virtual production is another important area for future development. The integration of real-time rendering with physical filmmaking environments can enable filmmakers to visualise digital environments and effects during production. Neural rendering and related technologies may contribute to the development of increasingly realistic and interactive digital environments (Tewari et al., 2020).

However, technological development must be accompanied by responsible practices concerning copyright, labour, consent, transparency, bias, and accountability. Recent research on AI and filmmaking suggests that AI is increasingly being considered as an augmentative technology that can support rather than necessarily replace human creative activity (Azzarelli et al., 2025).Artificial Intelligence has emerged as an important technological component of contemporary CGI and VFX research and production. Its applications extend across facial animation, character creation, pose estimation, motion capture, environmental generation, image segmentation, compositing, and post-production. The development of deep-learning architectures, GANs, transformers, diffusion models, neural-rendering systems, and foundation models has considerably expanded the technical possibilities available to filmmakers and VFX professionals.

AI-assisted systems can reduce selected repetitive operations, accelerate visual experimentation, support automated visual analysis, and assist in the generation and manipulation of digital content (Goodfellow et al., 2014; Ho et al., 2020; Rombach et al., 2022). At the same time, these technologies introduce significant questions concerning copyright, authorship, employment, performer consent, identity, transparency, and artistic control (Oberting, 2024; Mora & Llamas, 2025).

The development of AI in CGI and VFX should therefore be understood within a collaborative production framework. Computational systems can provide automation, analysis, and generative capabilities, while human professionals remain responsible for artistic direction, interpretation, evaluation, cultural sensitivity, and final creative decisions. The future of AI-assisted filmmaking will consequently depend not only on technological advancement but also on the development of appropriate professional, legal, and ethical practices.

REFERENCES

[1] Azzarelli, A., Anantrasirichai, N., & Bull, D. R. 2025. Intelligent cinematography: A review of AI research for cinematographic production. *Artificial Intelligence Review*, 58: 108.

[2] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. 2021. On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*: 610–623.

[3] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. 2021. *On the opportunities and risks of foundation models*. arXiv.

[4] Brock, A., Donahue, J., & Simonyan, K. 2019. Large scale GAN training for high fidelity natural image synthesis. *International Conference on Learning Representations*.

[5] Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. 2017. Realtime multi-person 2D pose estimation using part affinity fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 7291–7299.

[6] Chen, Y., Tian, Y., & He, M. 2020. Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, 192: 102897.

[7] Dang, Q., Yin, J., Wang, B., & Zheng, W. 2019. Deep learning based 2D human pose estimation: A survey. *Tsinghua Science and Technology*, 24(6): 663–676.

[8] Dhariwal, P., & Nichol, A. 2021. Diffusion models beat GANs on image synthesis. *Advances in Neural Information Processing Systems*, 34: 8780–8794.

[9] Dixon, W. W. 2019. *Synthetic Cinema: The 21st-Century Movie Machine*. Palgrave Macmillan.

[10] Fan, Y., Lin, Z., Saito, J., Wang, W., & Komura, T. 2022. FaceFormer: Speech-driven 3D facial animation with transformers. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 18770–18780.

[11] Feng, Y., Feng, H., Black, M. J., & Bolkart, T. 2021. Learning an animatable detailed 3D face model from in-the-wild images. *ACM Transactions on Graphics*, 40(4).

[12] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.

[13] Güler, R. A., Neverova, N., & Kokkinos, I. 2018. DensePose: Dense human pose estimation in the wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 7297–7306.

[14] He, K., Zhang, X., Ren, S., & Sun, J. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 770–778.

[15] Ho, J., Jain, A., & Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33: 6840–6851.

[16] Karras, T., Laine, S., & Aila, T. 2019. A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 4401–4410.

[17] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. 2020. Analyzing and improving the image quality of StyleGAN. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 8110–8119.

[18] Kerbl, B., Kopanas, G., Leimkuehler, T., & Drettakis, G. 2023. 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4).

[19] Kim, H., Garrido, P., Tewari, A., Xu, W., Thies, J., Nießner, M., Pérez, P., Richardt, C., Zollhöfer, M., & Theobalt, C. 2018. Deep video portraits. *ACM Transactions on Graphics*, 37(4).

[20] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. 2023. Segment anything. *Proceedings of the IEEE/CVF International Conference on Computer Vision*: 4015–4026.

[21] Long, J., Shelhamer, E., & Darrell, T. 2015. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 3431–3440.

[22] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. 2020. NeRF: Representing scenes as neural radiance fields for view synthesis. *European Conference on Computer Vision*: 405–421.

[23] Mora, R., & Llamas, M. 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.

[24] Oberting, V. A., IV. 2024. Generative artificial intelligence and copyright in the film and media industry. *Washington and Lee Law Review Online*, 82: 123.

[25] Pavllo, D., Feichtenhofer, C., Grangier, D., & Auli, M. 2019. 3D human pose estimation in video with temporal convolutions and semi-supervised training. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 7753–7762.

[26] Prince, S. 2012. *Digital Visual Effects in Cinema: The Seduction of Reality*. Rutgers University Press.

[27] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. 2021. Learning transferable visual models from natural language supervision. *Proceedings of the International Conference on Machine Learning*: 8748–8763.

[28] Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. 2021. Zero-shot text-to-image generation. *Proceedings of the International Conference on Machine Learning*: 8821–8831.

[29] Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K. V., Carion, N., Wu, C.-Y., Girshick, R., Dollár, P., & Feichtenhofer, C. 2024. *SAM 2: Segment Anything in Images and Videos*. arXiv.

[30] Ren, S., He, K., Girshick, R., & Sun, J. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28.

[31] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. 2022. High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 10684–10695.

[32] Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, K., Gontijo Lopes, R., Salimans, T., Ho, J., Fleet, D. J., & Norouzi, M. 2022. Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, 35.

[33] Sun, K., Xiao, B., Liu, D., & Wang, J. 2019. Deep high-resolution representation learning for human pose estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 5693–5703.

[34] Tewari, A., Fried, O., Thies, J., Sitzmann, V., Lombardi, S., Sunkavalli, K., Martin-Brualla, R., Simon, T., Saragih, J., Nießner, M., Pandey, R., Fanello, S., Wiles, O., & others. 2020. State of the art on neural rendering. *Computer Graphics Forum*, 39(2): 701–727.

[35] Thies, J., Zollhöfer, M., Stamminger, M., Theobalt, C., & Nießner, M. 2016. Face2Face: Real-time face capture and reenactment of RGB videos. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 2387–2395.

[36] Trigka, M., & Dritsas, E. 2025. A comprehensive survey of deep learning approaches in image processing. *Sensors*, 25(2): 531.

[37] Tsiavos, V., & Kitsios, F. 2025. The digital transformation of the film industry: How artificial intelligence is changing the seventh art. *Telecommunications Policy*, 49(8): 103021.

[38] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

[39] Xiao, B., Wu, H., & Wei, Y. 2018. Simple baselines for human pose estimation and tracking. *European Conference on Computer Vision*: 466–481.

[40] Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M.-H. 2024. Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys*, 56(4): Article 105.

[41] Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. 2024. A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57: 99.

[42] Zheng, C., Wu, W., Chen, C., Yang, T., Zhu, S., Shen, J., Kehtarnavaz, N., & Shah, M. 2021. Deep learning-based human pose estimation: A survey. *ACM Computing Surveys*, 54(1).