Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 4, 2222-2234
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i4.6539

Automated deepfake generation process based on AI painting technology

Qianru Liu[1], Ki-Hong Kim[2*]
[1,2]Department of Visual Contents, Dongseo University, Korea; 675liuqianru@gmail.com (K.H.K.)

Abstract: This paper addresses the significant data acquisition challenges in Deepfake technology by proposing an innovative methodology leveraging AI painting generation for creating facial datasets. The study aims to overcome traditional Deepfake limitations, including privacy concerns, copyright issues, and high costs associated with acquiring authentic portrait data. The research integrates AI art generation software (Midjourney with InsightFaceSwap) with Deepfake production software (DeepFace Lab) to create an automated workflow that generates diverse, consistent facial datasets for model training. Results demonstrate that AI-generated facial datasets can effectively substitute authentic human images while maintaining high-quality outputs. The proposed workflow significantly reduces data acquisition bottlenecks, mitigates legal risks, and substantially lowers production costs. Combining AI art generation with Deepfake technology offers a promising direction for advancing synthetic media creation while balancing innovation with ethical considerations. This methodology enhances the ethical application of Deepfake technology across entertainment, education, and digital media production, potentially transforming how synthetic visual content is created and used.
Keywords: *Artificial intelligence painting, Deepfake, Model training, Workflow.*

1. Introduction

The term "Deepfake" entered public consciousness 2017 when a Reddit moderator created a subreddit dedicated to videos using facial replacement technology. Since then, Deepfake technology has been used to develop various synthetic media, from celebrity impersonations to fabricated political speeches. While Deepfake technology offers extensive potential applications in entertainment and virtual reality, its development faces significant challenges. The primary obstacle is the need for substantial quantities of authentic human facial data for training. This requirement presents numerous privacy, copyright, and cost issues, creating a significant bottleneck in the advancement of Deepfake technology.

Concurrently, recent years have seen rapid progress in AI-generated imagery. Models such as Midjourney, Fooocus, and Stable Diffusion can produce highly realistic facial images. These AI image generation models can create unlimited quantities of diverse facial images on demand, potentially addressing the data constraints inherent in Deepfake model training.

This study aims to explore several key questions. Can AI-generated facial images serve as effective training data for Deepfake models? How does using AI-generated facial data impact the ethical and legal concerns associated with Deepfake technology? What are the potential cost and efficiency benefits of using AI-generated facial data in Deepfake production?

The researcher hypothesize that AI-generated facial images can provide high-quality, diverse training data for Deepfake models comparable to real facial images. The researcher also expects that expects using AI-generated facial data will significantly reduce legal and ethical risks associated with privacy infringement and intellectual property violations. Finally, we anticipate that utilizing AI-

History: Received: 26 February 2025; Revised: 15 April 2025; Accepted: 17 April 2025; Published: 24 April 2025
* Correspondence: 675liuqianru@gmail.com

generated imagery will substantially reduce the costs and increase the accessibility of Deepfake technology.

This study proposes an automated Deepfake generation process based on AI image generation technology. Unlike traditional methods that rely on real image data, our approach explores the use of AI-generated facial images exhibiting consistent appearances across various angles and expressions as a training dataset for Deepfake model development. By addressing these questions and testing these hypotheses, this study aims to contribute to developing more ethical, cost-effective, and accessible Deepfake technology, potentially broadening its applications across various fields, including entertainment, education, and digital media production.

2. Literature Review

2.1. Deepfake

Deepfake technology is a form of media manipulation that employs machine learning models such as VAE (Variational Autoencoders) and GAN (Generative Adversarial Networks) to merge and superimpose facial images or videos onto source material. By applying neural network techniques and large-scale sample learning, this technology synthesizes counterfeit content by integrating individual facial expressions and bodily movements into source images or videos [1]. Deepfake techniques can be categorized into three main types based on the manipulated facial regions and the intended purpose: identity swapping, facial reenactment, and attribute editing. Identity swapping involves replacing the identity of an individual in a source image with that of a target image while maintaining the background, effectively achieving a face-swapping effect. Utilizing deep learning algorithms, DeepFaceLab has proposed DF and LIAE structures [2] based on the Deepfake model, which currently represents the most natural and stable methods for identity swapping. Facial reenactment refers to identifying facial expressions without altering the individual's identity. This field has evolved through various approaches, including three-dimensional facial reconstruction techniques [3] the X2Face network [4] and occlusion-aware generators [5]. Recently, Hsu, et al. [6] introduced a three-dimensional facial keypoint detector with impressive visual results. Attribute editing involves adjusting specific facial appearance attributes such as skin tone, age, and hair. While early attempts at attribute editing by Perarnau, et al. [7] successfully modified targeted attributes, they struggled to maintain the stability of non-edited attributes. Subsequent research has explored various methods, including facial attribute classification [8] GRAF (generative radiance field) [9] and pi-GAN [10]. More recently, Xu, et al. [11] employed a Transformer-based framework to enhance the interaction between dual-space GANs, decoupling attributes' style and content representations, thereby improving the quality of generated images and the flexibility of attribute editing.

2.2. Artificial Intelligence Painting

Artificial Intelligence (AI) painting refers to automatically generating or creating visual artworks using AI algorithms and models through computational means. This technology transcends the limitations of traditional painting methods, offering new possibilities for artistic creation. AI painting techniques are primarily categorized into two modes: Text-to-Image and Image-to-Image. Text-to-Image technology generates corresponding image outputs based on textual input descriptions. Currently, the most widely utilized Text-to-Image models are diffusion models. These models simulate data by employing a diffusion process from data distribution to Gaussian distribution and a denoising process from Gaussian distribution back to data distribution. Chen, et al. [12] proposed Stable Diffusion, a Text-to-Image model capable of generating high-resolution images with rich details and semantic consistency. This model adopts the Latent Diffusion Model (LDM) paradigm for image generation. Unlike traditional diffusion models that operate in pixel space, LDM conducts diffusion in a lower-dimensional latent space. This approach significantly reduces the computational time required for training and inference while maintaining textual control over image content. Image-to-Image technology, conversely, utilizes AI algorithms to perform image transformation, restoration, or secondary creation based on existing source images, generating new images. The underlying principle involves Generative Adversarial Networks (GANs) learning the distribution of real images to generate convincing new ones. Although GANs have made significant advancements in promoting AI painting and image generation techniques, they exhibit limited control over output results when confronted with complex and diverse data, often producing random images. Furthermore, the images output by GANs tend to imitate existing works rather than innovate. It lacks human emotion and rationality, cannot truly understand and express the connotation and meaning of painting, and cannot reflect and respond to the social and cultural background of painting [13]. Despite the remarkable progress in AI painting technologies, several challenges persist. The control over specific image elements and the ability to generate truly novel content remain areas for improvement. Additionally, ethical considerations surrounding the use of training data and the potential for misuse of these technologies in creating deceptive content are ongoing concerns in the field. As AI painting continues to evolve, researchers focus on enhancing model interpretability, improving fine-grained control over generated images, and developing more sophisticated techniques for blending human creativity with AI capabilities.

2.3. Gaps in Existing Literature

While extensive research has been conducted on Deepfake techniques and AI painting technologies separately, there is a notable lack of studies that explore the integration of these two fields. Specifically, the potential of using AI-generated images as training data for Deepfake models remains largely unexplored. This gap presents an opportunity to address the ethical and legal concerns of using real human images in Deepfake production.

Furthermore, existing literature often focuses on improving the technical aspects of Deepfake and AI painting technologies, with less emphasis on developing workflows that could make these technologies more accessible and cost-effective for a broader range of applications. Our research aims to address these gaps by proposing an innovative workflow that combines AI art generation with Deepfake production, potentially opening new avenues for ethical and efficient synthetic media creation.

3. Comparison of the Results of Different Deepfake Compositing Software

3.1. Rope

As an emerging AI face-swapping tool, Rope is characterized primarily by its rapid processing speed and operational simplicity. Rope demonstrates exceptional performance in face-swapping completion time, attributable to its implementation of optimized algorithms and efficient parallel processing techniques. By leveraging GPU acceleration and multi-threaded processing, Rope can quickly accomplish complex face-swapping tasks. This high-efficiency processing capability renders Rope particularly suitable for applications requiring rapid generation of face-swapped content. Furthermore, Rope's user interface is designed with simplicity and intuitiveness, significantly lowering the barrier to entry for users. This design philosophy enhances user experience and expands the applicability of AI face-swapping technology. Non-professional users can quickly familiarize themselves with the tool and generate satisfactory face-swapping results without in-depth knowledge of the underlying technical principles. While emphasizing processing speed, Rope maintains a commendable output quality level by optimizing the model structure and training strategies. This approach enables Rope to significantly improve processing speed without compromising quality, achieving a favorable balance between efficiency and effectiveness in its algorithmic design.

However, Rope is not without limitations. To enhance processing speed, Rope reduces the resolution of source videos during the face-swapping process. While this approach improves efficiency, it presents a significant limitation in professional applications requiring high-definition output. Additionally, Rope's inability to train models for specific tasks results in suboptimal performance when processing complex scenarios, such as profile faces or faces under extreme lighting conditions.

DOI: 10.55214/25768484.v9i4.6539

Table 1.
Rope Performance Summary.

Ease of Use	Quality of Project Completion	Time Cost	Economic Cost	Data Requirement
2.0	3.5	1.0	0.0	1.0

***Note:** The evaluation is on a scale of 1-5 points.

3.2. Swapface

Swapface, as a mature AI face-swapping tool, primarily focuses on user-friendliness and technical reliability in its design philosophy. The platform offers a simple and intuitive user interface, enhancing user experience and broadening its user base. By simplifying operational processes and providing clear guidance, Swapface enables users without technical backgrounds to efficiently complete face-swapping tasks. This ease of use is crucial for the large-scale commercial application of AI technology. As a mature solution, Swapface maintains consistent performance across various scenarios. This stability is particularly important for commercial applications, as it ensures consistency and predictability of output results. Whether processing single photographs or extended video sequences, Swapface delivers commendable face-swapping effects. Like Rope, Swapface requires only a photograph to complete face-swapping, significantly streamlining user operations. This design improves efficiency and reduces the psychological burden on users. Users can obtain satisfactory face-swapping results without providing extensive training data or engaging in complex parameter adjustments. Furthermore, Swapface supports real-time preview functionality, allowing users to view effects before finalizing the replacement operation. This feature contributes to improving the accuracy of face-swapping results and user satisfaction. Users can make fine-tuning adjustments based on preview results to achieve outcomes more aligned with their expectations.

However, Swapface is not without limitations. Its business model requires users to subscribe to paid memberships for substantial face-swapping tasks. While this model ensures service quality and continuous technical updates, it restricts its application in certain high-frequency usage scenarios. It may incur high costs for users requiring large-scale face-swapping processing. Moreover, compared to some open-source solutions, Swapface offers limited customization capabilities. Users cannot deeply adjust algorithm parameters or train custom models, which may limit its flexibility in certain specialized application scenarios.

Table 2.
Swapface Performance Summary

Ease of Use	Quality of Project Completion	Time Cost	Economic Cost	Data Requirement
1.0	4.5	2.0	5.0	1.0

***Note:** The evaluation is on a scale of 1-5 points.

3.3. Faceswap

Faceswap, as a deep learning-based face-swapping tool, is characterized by its robust model training capabilities and user-friendly graphical interface. The tool excels in model training speed, rapidly achieving low loss values. This feature enhances model iteration efficiency and enables Faceswap to adapt quickly to new face replacement tasks. The efficient training process allows users to obtain satisfactory face-swapping models relatively quickly, which is particularly valuable for applications requiring frequent model updates. Furthermore, Faceswap provides an intuitive graphical user interface that simplifies the operational process. Users need only import source, target face videos, and select appropriate algorithms and parameters to generate face-swapping results. This design significantly lowers the technical barrier, making complex AI face-swapping technology accessible to a broader user base. The graphical interface offers basic operational functions and includes advanced features such as data preprocessing and model training progress visualization, providing users with comprehensive control capabilities. Faceswap supports multiple face-swapping algorithms, allowing users to select the most suitable algorithm based on specific requirements. This flexibility enables Faceswap to adapt to various application scenarios, from simple photo face-swapping to complex video face-swapping. As an open-source project, Faceswap benefits from the support of an active developer community, ensuring continuous software updates and improvements while providing users with rich learning resources and technical support.

However, Faceswap is not without limitations. Training models solely through source video materials may result in less refined generation results. This limitation may become apparent when handling high-quality face-swapping tasks, particularly in scenarios requiring precise capture of facial details and expressions. Additionally, Faceswap's installation environment is relatively complex, which may create usage barriers for some users, especially those without technical backgrounds. The complex environment configuration increases the difficulty of initial use and may also affect software stability and performance. This complexity may limit Faceswap's widespread adoption, particularly in commercial applications requiring rapid deployment.

Table 3.
Faceswap Performance Summary.

Ease of Use	Quality of Project Completion	Time Cost	Economic Cost	Data Requirement
5.0	4.0	3.0	0.0	3.0

***Note:** The evaluation is on a scale of 1-5 points.

3.4. Deepface Lab

Deepface Lab, as an open-source deepfake system, is renowned for its powerful customization capabilities and high-quality output results. As an open-source project, Deepface Lab reduces usage costs and provides a platform for community contributions and technological innovation. Developers and researchers can freely access and modify the source code, fostering rapid development and innovation in AI face-swapping technology. The active participation of the open-source community ensures that Deepface Lab continuously improves and adapts to new technological trends. Deepface Lab boasts robust customization capabilities, allowing users to train customized models based on specific requirements, significantly enhancing the quality and adaptability of face-swapping results. It offers a rich array of parameter settings and model selection options, enabling users to fine-tune various aspects of the face-swapping process. This high customizability allows Deepface Lab to adapt to diverse and complex face-swapping needs, distinguishing it in professional applications. The system can generate high-fidelity face-swapping results suitable for long-duration video processing and cinema-grade face-swapping requirements[2]. Through advanced deep learning algorithms and sophisticated facial feature extraction techniques, Deepface Lab precisely captures and reproduces facial expressions, lighting changes, and subtle movements. This high-quality output gives Deepface Lab a significant advantage in professional visual effects production. Furthermore, Deepface Lab provides a comprehensive workflow encompassing multiple stages, including data preprocessing, model training, face-swapping generation, and post-processing. This holistic functional design allows users to complete entire face-swapping projects on a unified platform, improving work efficiency and result consistency.

However, Deepface Lab's high-quality output comes at the cost of high data requirements. It necessitates a large amount of material for model training, which not only increases usage costs but may also involve legal issues related to data privacy and copyright. In scenarios with limited resources, these high data requirements may somewhat restrict Deepface Lab's application.

Table 4.
Deepface Lab Performance Summary.

Ease of Use	Quality of Project Completion	Time Cost	Economic Cost	Data Requirement
3.5	5.0	4.0	0.0	5.0

***Note:** The evaluation is on a scale of 1-5 points.

In conclusion, Rope's performance in complex scenarios and output quality still has room for improvement. The economic costs associated with Swapface's business model may limit its application in certain high-frequency usage scenarios. The complexity of Faceswap's installation environment and issues with result refinement may affect its application in certain professional scenarios. So, the author considers Deepface Lab the optimal choice for the current workflow. This assessment is based on its superior output quality, extensive customization options, and suitability for professional-grade applications.

Table 5.
Performance comparison of various Deepfake software.

Software	Ease of Use	Quality of Project Completion	Time Cost	Economic Cost	Data Requirement
Rope	2.0	3.5	1.0	0.0	1.0
Swapface	1.0	4.5	2.0	5.0	1.0
Faceswap	5.0	4.0	3.0	0.0	3.0
Deepface Lab	3.5	5.0	4.0	0.0	5.0

***Note:** The evaluation is on a scale of 1-5 points.

4. Comparison of the Effect of Different AI Drawing Software in Generating the Face Material Needed for Deepfake

4.1. Stable Diffusion

As one of the leading AI image generation models, Stable Diffusion has been widely applied to various image generation tasks. However, testing has revealed significant deficiencies when directly utilizing the Image-to-Image mode to generate human faces. As shown in Figure 1 [14] the facial expressions in the generated images often have insufficient range or are uncontrollable, to simulate the rich spectrum of human facial expressions authentically. These issues render AI-generated facial images produced in this manner inadequate for meeting the high-quality requirements necessary for Deepfake training.

Figure 1.
Rigidity of Expression [14].

To overcome this challenge, a currently viable approach is to utilize the local redrawing (Inpaint) function within the image-to-image mode, obtaining the required facial image materials through face swapping. This method primarily relies on a model called "IP-Adapter Face ID Plus V2," which can precisely control the image generation process under Stable Diffusion's ControlNet architecture. As shown in Figure 2, while this method can, to some extent, address the issues of rigid expressions and facial instability, it often fails to achieve a natural transition when replacing faces with significantly different skin tones. This results in output images that appear artificial and mask-like, falling short of the required quality level for training.

Figure 2.
Mask-Like due to Inconsistent Skin Tone [14].

In the context of Deepfake technology, these limitations of Stable Diffusion present significant challenges. The lack of diverse and natural facial expressions in the generated images could lead to Deepfake models that struggle to reproduce a wide range of human emotions and expressions accurately. Furthermore, the issues with skin tone transitions could result in unrealistic or inconsistent appearances in the final Deepfake outputs, particularly when applied to individuals with diverse ethnic backgrounds.

4.2. Midjourney

Midjourney, a mainstream AI painting software, provides users with an efficient and convenient creative platform. Currently, there are two methods to generate a series of facial images with consistent features using Midjourney: adding command parameters and using plugins. By utilizing Midjourney's character feature control parameter "--cref" and source image style inheritance parameter "--sref", users can generate facial images with realistic and vivid expressions while ensuring consistency in prominent facial features. However, even with these methods, the series of images generated by Midjourney may still inevitably contain subtle differences. As illustrated in Figure 3 [14] within images of the same girl, details such as the density and distribution of freckles on the face and eye color can vary to some degree. When creating a series of portraits, these differences tend to be amplified as the number of generated images increases, resulting in decreased uniformity as more images are produced.

Figure 3.
Creating Faces by Adding Commands [14].

To address this issue, a plugin called InsightFaceSwap presents a more effective alternative. Users only need to add the InsightFaceSwap bot to their server to input a source face, save it as a Face ID, and proceed with face swapping. Leveraging advanced algorithms, the InsightFaceSwap plugin demonstrates excellent performance in terms of both uniformity and stability, significantly enhancing the consistency and reliability of generated images.

The integration of Midjourney with Deepfake technology offers several advantages. The high degree of control over facial features and expressions provided by Midjourney's command parameters allows for creating diverse yet consistent training datasets. This is particularly valuable for Deepfake models, as it enables the generation of a wide range of facial expressions and angles while maintaining the core identity of the subject. As shown in Figure 4, the InsightFaceSwap plugin further enhances this capability by ensuring high consistency across multiple generated images, which is essential for training robust Deepfake models.

Figure 4.
Creating Faces with InsightFaceSwap

Moreover, the ability to generate diverse facial images without relying on real human subjects addresses some ethical concerns associated with Deepfake technology, such as privacy infringement and

DOI: 10.55214/25768484.v9i4.6539

consent issues. Using AI-generated faces as the source material for Deepfake training, researchers and developers can explore and advance the technology while mitigating some potential negative societal impacts.

However, it's important to note that integrating AI-generated images into Deepfake workflows is not without challenges. The subtle inconsistencies in Midjourney-generated images, such as freckle patterns or eye color variations, could potentially lead to artifacts or inconsistencies in the final Deepfake output. Careful quality control and additional post-processing steps may be necessary to ensure the highest quality results.

In conclusion, the most viable solution is employing Midjourney with the InsightFaceSwap plugin to produce the facial materials required for Deepfake training. This approach organically integrates AI painting models with facial recognition technology, fully exploiting both advantages. It generates facial images with highly consistent features and ensures a rich diversity of expressions while maintaining natural perspective angles and a high degree of realism. These characteristics make it particularly well-suited for creating high-quality training datasets for Deepfake models, potentially advancing the field while addressing some of its ethical challenges [14].

5. Workflow

Deepfake video production techniques are garnering increasing attention and application in the rapid development of artificial intelligence technology. However, traditional Deepfake video production methods typically require a substantial quantity of authentic human facial images as training data, which is costly and fraught with legal risks pertaining to privacy and copyright issues. Consequently, this paper will elucidate an innovative workflow that optimizes the existing Deepfake video production process by utilizing the AI art generation software Midjourney and the Deepfake production software Deepface Lab. As shown in Figure 5, the core logic of this workflow revolves around employing AI art generation software to create the Source (Src) materials required for training Deepface Lab models, thereby supplanting the need for Src materials obtained through photographing real individuals. This approach significantly reduces production costs and simplifies the production process. Furthermore, since all materials used are AI-generated virtual characters, it circumvents any issues related to real individuals' publicity rights. Consequently, the Deepfake models trained through this method will have broader applications across various scenarios. This innovative methodology addresses several critical issues in Deepfake technology, including cost reduction, simplification of processes, and mitigation of legal risks. By leveraging the capabilities of AI art generation, it opens up new possibilities for the ethical and efficient production of Deepfake content, potentially revolutionizing the landscape of digital media creation and manipulation.

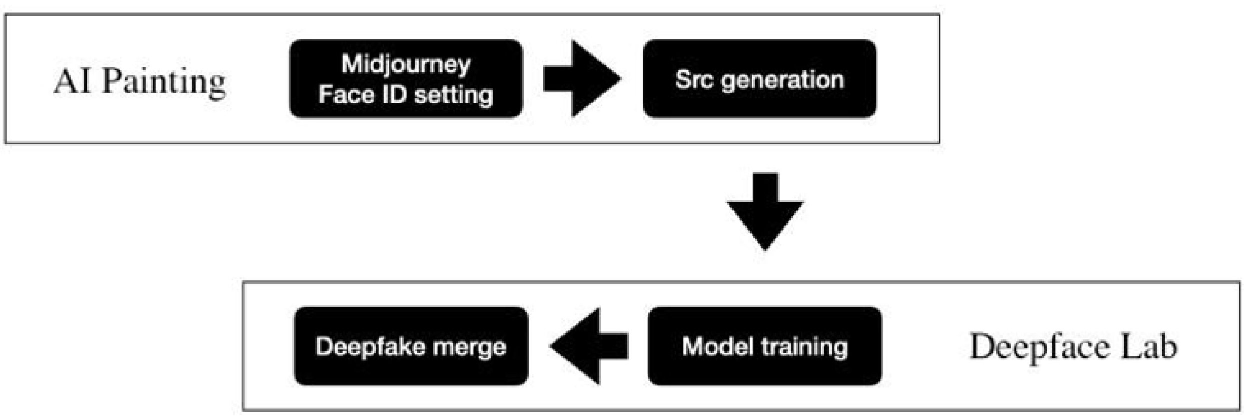

Figure 5.
Workflow Design.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 4: 2222-2234, 2025
DOI: 10.55214/25768484.v9i4.6539

As shown in Figure 6, the proposed workflow for creating high-quality Deepfake videos using AI-generated images involves several interconnected steps. First, the researcher generates a photorealistic facial image using Midjourney as the source face. This image is then saved as a Face ID within Midjourney's InsightFaceSwap plugin. Concurrently, the researcher uses Metahuman or similar three-dimensional character creation software to render facial close-up images encompassing various angles and expressions, which serve as target images for face swapping and training materials. The data preprocessing stage is critical for ensuring the quality and consistency of the training data. The researcher imports the generated images into DeepFace Lab software, which employs advanced facial detection and alignment algorithms. These algorithms identify key facial landmarks and normalize the images for orientation, size, and position. This normalization process is essential for reducing variability in the training data that could negatively impact model performance.

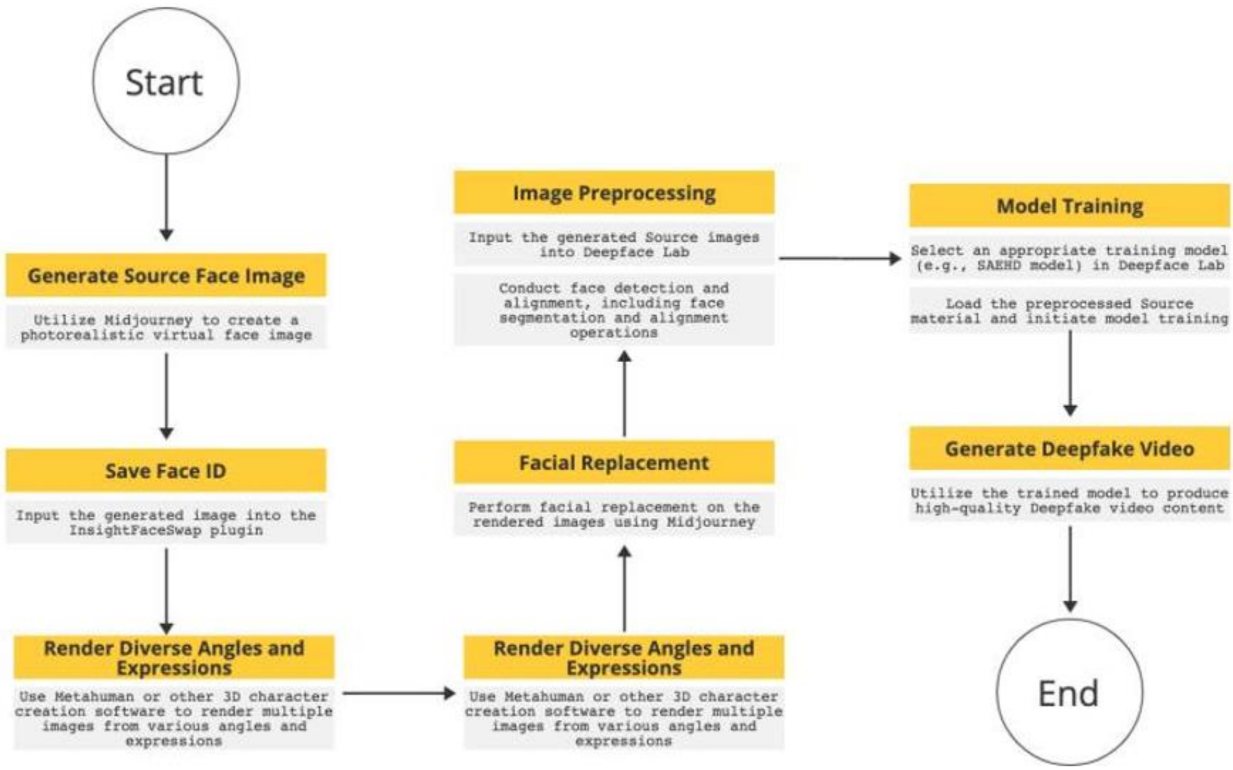

Figure 6.
Workflow.

An appropriate training model within DeepFace Lab is selected based on several factors. The researcher considers the complexity of the facial features, the diversity of expressions and angles in the dataset, and the desired output quality. For datasets with high variability in expressions and angles, the researcher opts for models with greater capacity for learning complex representations, such as those based on deeper neural network architectures. Conversely, the researcher chooses simpler models for more homogeneous datasets to avoid overfitting. Once the model is selected and the data is preprocessed, the researcher initiates the training using the prepared source materials. This training process involves iterative optimization of the model parameters to minimize the difference between the generated and target faces. The researcher monitors key metrics such as loss values and the visual quality of interim outputs to gauge training progress and make necessary adjustments. Upon completion of training, the resultant model can be employed to swiftly produce high-quality Deepfake video content.

This approach synergistically combines AI art generation and Deepfake technology to create a streamlined, cost-effective, and legally compliant workflow. It mitigates privacy and image rights concerns while maintaining a high quality and realism standard in the final output. Consequently, this methodology has the potential to significantly broaden the applications of Deepfake technology across various fields, including entertainment, education, and digital media production.

6. Conclusions

Deepfake video production has emerged as an increasingly prominent and intriguing field in the rapidly evolving landscape of artificial intelligence technology. This technology, which employs AI-driven facial modifications to merge source and target facial features in videos seamlessly, has ushered in unprecedented creativity and novel experiences in video production. However, traditional Deepfake video creation methods are fraught with numerous drawbacks, such as the requirement for extensive real human facial imagery as training data, which not only incurs substantial costs but also poses legal risks related to privacy breaches and infringement of image rights. The advent of AI-generated art technologies has provided a unique approach to addressing these challenges.

This paper proposes an innovative workflow that integrates the AI art generation software Midjourney with the Deepfake production software Deepface Lab, alongside Metahuman rendering imagery, to create Deepfake videos. This methodology organically combines AI art generation, facial recognition technology, and Deepfake video production techniques, leveraging the strengths of various AI technologies. It serves as a vivid example of promoting cross-disciplinary integration and innovative applications of AI technology.

The proposed workflow addresses several challenges in traditional Deepfake video production. By utilizing AI art generation technology to create source materials (Src), the need to hire human models or conduct expensive photoshoots is eliminated, significantly reducing production costs. The AI art software Midjourney, equipped with the InsightFaceSwap plugin, generates Src materials with highly consistent facial features while offering diverse expressions and angles, meeting training requirements for various scenarios and needs. The workflow is highly automated, requiring operators to provide only necessary textual descriptions and control parameters, enhancing production efficiency. Moreover, since the image materials generated by AI art are virtual character representations, they do not involve any real human image rights issues, reducing potential legal risks. These advantages position the proposed methodology as a promising approach for the ethical and efficient production of Deepfake content, with potential applications across multiple domains, including video production, IP creation, virtual reality, and the metaverse.

This innovative approach addresses the practical challenges in Deepfake video production and exemplifies the synergistic potential of integrating various AI technologies. By leveraging the strengths of AI art generation, facial recognition, and Deepfake techniques, it opens up new possibilities for creative content production while mitigating ethical and legal concerns. As such, this methodology represents a significant step forward in the evolution of AI-driven media creation, offering a glimpse into the future of digital content production and manipulation.

However, this study has several limitations that should be acknowledged. While the proposed method generates high-quality Deepfake videos, a systematic comparison with traditional methods using real human images as training data was not conducted. Although using AI-generated images mitigates some ethical concerns, it doesn't eliminate all ethical issues surrounding Deepfake technology. The long-term stability and consistency of AI-generated facial datasets in Deepfake model training have not been thoroughly tested. Additionally, while the potential for broad application is discussed, the study primarily focused on video production.

Future research directions include conducting rigorous comparisons between Deepfake models trained on AI-generated images versus those trained on real human images, focusing on quality, efficiency, and ethical considerations. Developing a comprehensive ethical framework for using AI-generated imagery in Deepfake production, and addressing potential misuse and societal impacts is another important avenue. Exploring the application of this methodology in fields beyond video production, such as virtual reality, augmented reality, and interactive media, could yield valuable insights. Investigating ways further to improve the quality and diversity of AI-generated facial datasets, incorporating more advanced AI models as they become available, is also a promising direction. Lastly, conducting user perception studies to gauge public reception and potential concerns regarding Deepfake content created using AI-generated imagery could provide valuable feedback for future developments.

Future studies can build upon this work by addressing these limitations and pursuing these research directions further to advance the ethical and efficient Deepfake production field. This innovative approach offers solutions to current challenges in Deepfake technology and opens up new possibilities for AI's creative and responsible use in media production.

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Acknowledgement:

This research was supported by BK21 Four Future Visual Contents for Creative Convergence Education Research Group in Dongseo University.

Copyright:

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Copyright:

© 2025 by the authors. This open-access article is distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

References

[1] W.-B. Zhou, W.-M. Zhang, N.-H. Yu, H. Zhao, H. Liu, and T. Wei, "An overview of deepfake forgery and defense techniques," *Journal of Signal Processing*, vol. 37, no. 12, pp. 2338–2355, 2021. https://doi.org/10.16798/j.issn.1003-0530.2021.12.007

[2] K. Liu, I. Perov, D. Gao, N. Chervoniy, W. Zhou, and W. Zhang, "Deepfacelab: Integrated, flexible and extensible face-swapping framework," *Pattern Recognition*, vol. 141, p. 109628, 2023. https://doi.org/10.1016/j.patcog.2023.109628

[3] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt, and M. Nießner, "Face2face: Real-time face capture and reenactment of rgb videos," in *Proceedings of the IEEE cConference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2016.262*, 2016, pp. 2387-2395.

[4] O. Wiles, A. Koepke, and A. Zisserman, "X2face: A network for controlling face generation using images, audio, and pose codes," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 670-686.

[5] A. Siarohin, S. Lathuilière, S. Tulyakov, E. Ricci, and N. Sebe, "First order motion model for image animation," *Advances in Neural Information Processing Systems*, vol. 32, 2019. https://doi.org/10.48550/arXiv.2003.00196

[6] G.-S. Hsu, C.-H. Tsai, and H.-Y. Wu, "Dual-generator face reenactment," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 642-650.

[7] G. Perarnau, J. Van De Weijer, B. Raducanu, and J. M. Álvarez, "Invertible conditional gans for image editing," *arXiv preprint arXiv:1611.06355*, 2016. https://doi.org/10.48550/arXiv.1611.06355

[8] Z. He, W. Zuo, M. Kan, S. Shan, and X. Chen, "Attgan: Facial attribute editing by only changing what you want," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5464-5478, 2019. https://doi.org/10.1109/TIP.2019.2917986
[9] K. Schwarz, Y. Liao, M. Niemeyer, and A. Geiger, "Graf: Generative radiance fields for 3d-aware image synthesis," *Advances in Neural Information Processing Systems*, vol. 33, pp. 20154-20166, 2020. https://doi.org/10.5555/3495724.3497416
[10] E. R. Chan, M. Monteiro, P. Kellnhofer, J. Wu, and G. Wetzstein, "Pi-gan: Periodic implicit generative adversarial networks for 3d-aware image synthesis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 5799-5809.
[11] Y. Xu *et al.*, "Transeditor: Transformer-based dual-space gan for highly controllable facial editing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 7683-7692.
[12] L. Chen, Z. Li, R. K. Maddox, Z. Duan, and C. Xu, "Lip movements generation at a glance," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 520-535.
[13] J. Li, J. Zhong, S. Liu, and X. Fan, "Opportunities and challenges in AI painting: The game between artificial intelligence and humanity.," *Journal of Big Data and Computing*, vol. 2, no. 1, pp. 1-15, 2024. https://doi.org/10.1007/978-3-030-01234-2_32
[14] Q.-R. Liu and K.-H. Kim, "A comparative study of uniformity and stability of real faces generated by AI painting software," in *Proceedings of the General Assembly of the Korea Content Association*, 2023, pp. 435–436.