International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 14, Issue 1, 2026

Faceforge: Realistic Human Face Generation Using Gan's

Saleha Farha[1], Charitha Reddy Ravula[2], Nandini Kummari[3], Pasupunoor Nishita[4]
[1]Assistant Professor; Department Of Information Technology Bhoj Reddy Engineering College For Women Hyderabad, India.
[2,3,4]B.Tech Students; Department Of Information Technology Bhoj Reddy Engineering College For Women Hyderabad, India.
Mail Id; reddyrcharitha@gmail.com[2] nandukummari369@gmail.com[3], pasupunoornishita@gmail.com[4]

Abstract
The synthesis of photorealistic human faces using Generative Adversarial Networks (GANs) has gained significant attention due to its capability to produce visually convincing and diverse images. This work aims to enhance facial image generation quality by addressing limitations present in conventional GAN-based approaches and refining both the generator and discriminator modules of a Deep Convolutional Generative Adversarial Network (DCGAN). The proposed methodology focuses on improving the discriminator's capacity to effectively differentiate between authentic and synthesized facial images while simultaneously optimizing the generator to create high-resolution, artifact-free outputs with improved visual realism.To achieve this objective, large-scale and diverse facial datasets, including CelebA, are utilized to train the model and improve generalization. A comprehensive evaluation of existing GAN architectures is conducted, followed by the implementation of modified DCGAN variants incorporating architectural refinements and optimization strategies. These enhancements contribute to improved image fidelity, better texture representation, and enhanced structural consistency. Experimental results indicate noticeable improvements in generated facial images, with outputs closely resembling real-world photographs while minimizing common artifacts such as blurring, asymmetry, and unnatural visual patterns. The proposed approach demonstrates the effectiveness of refined DCGAN architectures in producing high-quality photorealistic human faces suitable for various computer vision applications.

Keywords: *Generative Adversarial Networks, DCGAN, Photorealistic Face Generation, Deep Learning, Image Synthesis, Generator–Discriminator Architecture, Facial Image Generation, Computer Vision.*

Introduction

Artificial Intelligence has witnessed rapid advancement in recent years, particularly in generative modeling techniques capable of producing realistic synthetic data. Among these approaches, Generative Adversarial Networks (GANs) have become one of the most influential frameworks for image synthesis. A notable variant, the Deep Convolutional Generative Adversarial Network (DCGAN), integrates convolutional neural networks into the GAN architecture, enabling the generation of high-quality visual content. Since its introduction, DCGAN has been widely applied in various domains including medical imaging, computer vision, and multimedia applications, demonstrating strong performance in learning complex visual patterns.The generation of photorealistic human faces has emerged as an important research direction due to its potential applications in entertainment, virtual reality, biometric systems, digital avatars, and privacy-preserving technologies. Producing realistic facial images requires capturing fine-grained details such as texture, symmetry, illumination, and facial structure. However, achieving high-quality outputs remains challenging due to issues such as training instability, mode collapse, and visual artifacts commonly observed in GAN-based systems.This work investigates advanced techniques for photorealistic face generation using a refined DCGAN architecture. The proposed approach begins with the development of a baseline DCGAN model, which is subsequently enhanced through architectural modifications including additional convolutional layers, optimized dense layers, and improved training strategies. A large-scale facial dataset is employed to ensure diversity and robustness in the generated images. The adversarial training process enables the generator to progressively improve its ability to synthesize realistic faces, while the discriminator simultaneously learns to differentiate authentic images from generated samples.The generated outputs are evaluated using qualitative and quantitative measures to assess realism, clarity, and structural consistency. The customized architecture is designed to reduce common GAN limitations such as mode collapse and image distortions while improving texture accuracy and facial symmetry. Overall, the proposed system aims to enhance generative modeling performance and produce highly convincing human face images suitable for practical applications.

Purpose of the Project

The primary objective of this project is to design and implement a deep learning framework capable of generating highly realistic human facial images using a DCGAN architecture. The system aims to improve the generator's ability to synthesize photorealistic faces while strengthening the discriminator's performance in distinguishing real images from synthetic ones. Another key goal is to address common challenges associated with GANs, including image blurring, visual artifacts, and mode collapse. By training the model on a diverse facial dataset, the system learns varied facial attributes and enhances generalization. The project also focuses on experimenting with customized architectural modifications to improve image sharpness, texture consistency, and structural realism. Furthermore, the generated images are evaluated using suitable performance metrics to measure their visual quality and authenticity.

Existing System

Traditional image generation approaches such as Variational Autoencoders (VAEs), PixelCNN, and diffusion-based models have been used for synthetic image creation. However, these techniques have certain limitations when applied to photorealistic face generation. VAEs often produce smooth but blurry images due to their probabilistic reconstruction process. PixelCNN models are capable of generating detailed outputs but suffer from high computational complexity and slow sampling speed. Diffusion models have demonstrated strong performance in image synthesis, yet they require multiple iterative sampling steps, making them computationally expensive. Although GAN-based models generate sharper images compared to these approaches, they are prone to issues such as unstable training, mode collapse, and limited diversity in outputs. These challenges highlight the need for an improved architecture that can efficiently generate high-quality and realistic human facial images.

Proposed System

The proposed system utilizes a Deep Convolutional Generative Adversarial Network to generate diverse and photorealistic human face images. The architecture consists of a generator that produces synthetic facial images and a discriminator that evaluates their authenticity. Both networks are trained simultaneously in an adversarial manner, allowing continuous improvement in image quality. The training process uses a large-scale facial dataset that is preprocessed through resizing, normalization, and augmentation techniques to enhance learning efficiency. To improve stability and reduce common GAN-related issues, techniques such as batch normalization, label smoothing, and dropout are incorporated into the architecture. These enhancements help mitigate mode collapse, improve convergence, and enhance output diversity. The proposed system aims to generate high-quality, visually consistent, and unbiased facial images suitable for real-world artificial intelligence applications.

Survey

A comprehensive review of existing generative adversarial network (GAN)–based systems was conducted to design an effective AI-driven image generation and deep fake detection framework. The purpose of this survey is to examine recent advancements in generative models, analyze deep learning–based detection strategies, and identify research gaps that motivate the proposed approach. Early work on deep convolutional generative adversarial networks demonstrated that convolutional architectures significantly improve image synthesis quality and training stability. This framework showed the capability to generate structured visual data and control image attributes, making it a foundational model for many later developments.Deep fake detection techniques have also evolved alongside generative models. Several studies combined error-level analysis with convolutional neural networks to enhance the detection of manipulated images. These approaches improved classification accuracy by identifying inconsistencies introduced during image synthesis. Convolutional neural networks have further been applied in domains such as medical imaging, where they achieved strong performance in tasks like disease classification, highlighting the robustness of deep learning for visual analysis. Similar concepts have been adapted for face forgery detection, where deep residual networks demonstrated improved capability in distinguishing authentic images from manipulated ones.Research has also explored different strategies for managing facial identity manipulation, including appearance-based, feature-based, model-based, and hybrid methods. These techniques aim to handle challenges such as variations in pose, lighting, expression, and background clutter. Effective detection often relies on combining both content-level features and trace-level artifacts; however, subtle manipulations remain difficult to identify. To address high-dimensional data challenges, sparse-based classifiers and fast optimization techniques have been introduced. Methods such as orthogonal matching pursuit and support vector machines improve classification efficiency and performance, particularly when working with imbalanced datasets.Several deep learning–based face recognition and detection models, including local binary pattern histograms and Fisherface-based approaches, have been enhanced using deep belief networks to capture both structural and statistical features. Meanwhile, various GAN variants have been proposed to improve image synthesis quality and training stability. Techniques such as least-squares GANs and blending-based architectures have demonstrated improved realism and reduced artifacts, particularly when trained on large-scale facial datasets. Additional research has focused on identifying synthetic fingerprints left by GAN-generated images, enabling more accurate detection of manipulated content. Synthetic data generated using GANs has also been applied in other domains, such as agriculture, where artificially generated datasets improved classification accuracy for plant disease detection tasks.More recent studies introduced multi-channel GAN architectures combined with multi-scale residual networks for segmentation tasks in medical imaging. These models achieved accurate extraction of complex structures such as retinal vessels and optic discs, supporting diagnostic applications. Beyond traditional GAN models, newer architectures have emerged that provide enhanced control over facial attributes and improved image fidelity. Advanced generative models enable the creation of realistic facial representations, three-dimensional avatars, and intrinsic feature extraction. Continuous improvements in these architectures highlight their versatility in image synthesis, face editing, and domain-specific applications.

Requirement Analysis

Functional Requirements

Functional requirements describe the specific operations that the proposed system must perform and define the relationship between system inputs and expected outputs. These requirements outline the workflows involved in each stage of the model pipeline and ensure that all core functionalities operate as intended. The system begins with data preprocessing, where facial images from the dataset are resized, normalized, and prepared for training. Following preprocessing, the model training phase involves configuring and training the DCGAN architecture through adversarial learning between the generator and discriminator networks.Once training is complete, the system performs image generation, where the trained generator produces synthetic facial images from random noise vectors. The generated outputs are then subjected to performance evaluation using appropriate qualitative and quantitative metrics to assess realism, clarity, and structural consistency. Finally, the system provides result visualization, enabling users to inspect generated images, monitor training progress, and analyze performance trends. These functional components collectively define the operational workflow of the proposed deep learning framework.

Non-Functional Requirements

Non-functional requirements specify the quality attributes and operational constraints of the system. The proposed platform must maintain high performance to ensure efficient training and image generation within reasonable computational time. Security considerations include safeguarding datasets and model outputs while ensuring secure access to development environments. Reliability is essential to guarantee consistent model behavior and stable training without frequent interruptions. The system should also provide usability through a simple and intuitive interface that supports easy experimentation and result interpretation.Scalability is required to accommodate larger datasets and more complex model architectures in future extensions. Maintainability is another important aspect, allowing developers to update model parameters, integrate new modules, and improve performance without major restructuring. Portability ensures that the system can operate across different platforms and computing environments with minimal configuration changes. Additionally, ethical compliance must be considered, ensuring responsible use of generated facial images and adherence to data privacy and fairness guidelines. These non-functional requirements collectively ensure robustness and long-term sustainability of the proposed system.

Software Requirements

The implementation of the proposed system requires a set of software tools and libraries to support model development and experimentation. The CelebA dataset is used as the primary source of facial images for training and evaluation. The model is implemented using the Python programming language due to its flexibility and extensive support for deep learning applications. The PyTorch framework is employed to build and train the DCGAN architecture. Supporting libraries such as NumPy are used for numerical operations, while Matplotlib assists in visualizing results. Image preprocessing is handled using the Python Imaging Library (PIL) and torchvision utilities. The development and experimentation are carried out using Google Colab, which provides access to GPU resources for faster training. Version control and collaboration are managed using GitHub to track modifications and facilitate teamwork.

Hardware Requirements

The system requires moderate computational hardware to support deep learning model training and image generation. A processor equivalent to Intel i5 or Intel i7 is recommended to ensure efficient execution of preprocessing and training operations. A minimum of 12 GB RAM is suggested to handle dataset loading and model computations without memory constraints. Storage capacity of approximately 100 GB is preferred to accommodate datasets, trained models, and generated outputs. The system can operate on Windows 10 or compatible operating systems, with optional GPU acceleration improving training speed and overall performance. These hardware specifications provide sufficient resources for implementing and evaluating the proposed DCGAN-based face generation framework.

Design

System Architecture

System architecture describes the overall structure and operational flow of the proposed solution,

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 14, Issue 1, 2026

highlighting how different components interact to achieve the desired objectives. It provides a high-level conceptual view of the system, typically illustrated through architectural diagrams, and serves as a blueprint for development. The architecture ensures that all modules work together cohesively while satisfying both functional and non-functional requirements.The proposed system begins with dataset acquisition, where facial images are collected from a large-scale dataset. These images undergo preprocessing steps such as resizing, normalization, and augmentation to prepare them for training. The processed data is then passed to the DCGAN model, which consists of generator and discriminator networks trained through adversarial learning. The generator produces synthetic facial images from random noise, while the discriminator evaluates their authenticity by comparing them with real images.

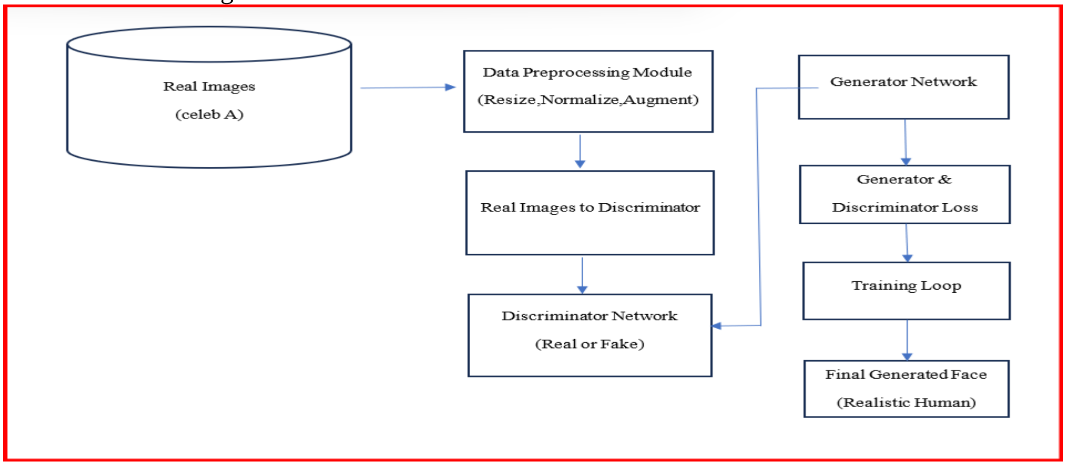


Fig. 1 System Architecture

During training, both networks are optimized iteratively to improve performance and generate realistic outputs. Once training is complete, the system produces synthetic facial images that are evaluated using performance metrics to assess visual quality and realism. The final stage includes visualization of generated images and analysis of results. This architecture ensures a structured workflow from data preparation to model evaluation, enabling efficient implementation and scalability for future enhancements.

Technical Architecture

The technical architecture of the proposed image generation system is designed to be modular, scalable, and efficient. The implementation environment is primarily based on Python, which supports deep learning development and integration of machine learning libraries. The training pipeline is built using the PyTorch framework, enabling the construction of the DCGAN model and efficient GPU-based computation. The system incorporates preprocessing modules responsible for image normalization, resizing, and dataset management before feeding the data into the model.The generator network transforms random noise vectors into synthetic facial images using transposed convolution layers, batch normalization, and activation functions. The discriminator network employs convolutional layers to extract features and classify images as real or generated. These components interact through adversarial training to continuously improve image quality. Supporting libraries such as NumPy and torchvision assist in numerical operations and dataset handling, while Matplotlib is used for visualizing generated results and training progress.

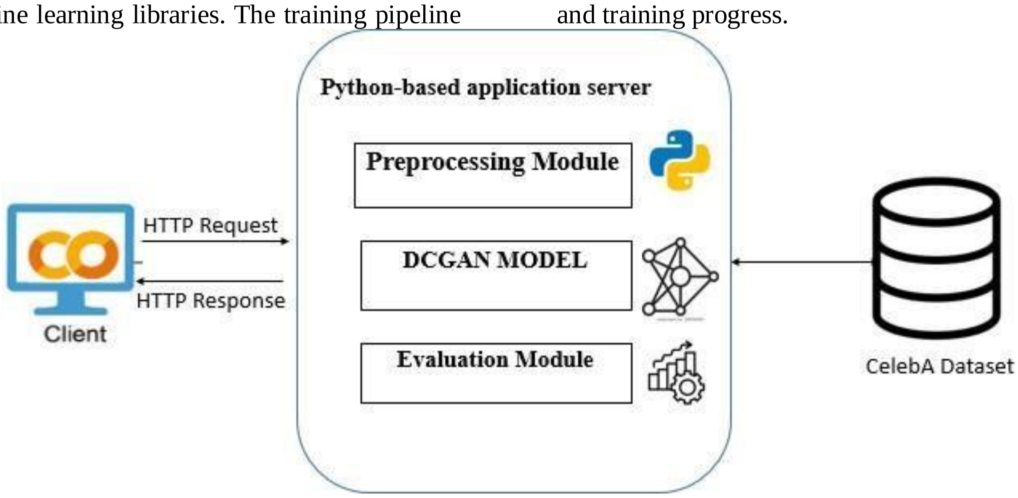


Fig 2 Technical Architecture

The development environment, such as Google Colab, provides computational resources including GPU acceleration for faster training. Model outputs, checkpoints, and generated images are stored for evaluation and future experimentation. The modular structure of the technical architecture allows easy modification of model parameters, integration of advanced GAN variants, and scalability for handling larger datasets. This design ensures efficient training, improved realism in generated images, and flexibility for further research and development.

Methodology

Generative Adversarial Networks (GANs) are deep learning models designed to generate new data samples by learning the underlying distribution of training data. A widely adopted variant, Deep Convolutional Generative Adversarial Network (DCGAN), is particularly effective for image synthesis tasks because convolutional layers capture spatial relationships within images. The proposed methodology employs a DCGAN architecture consisting of generator and discriminator networks trained in an adversarial manner. Activation functions such as ReLU and Tanh are used in the generator to enhance non-linearity and produce normalized outputs, while batch normalization stabilizes training and accelerates convergence.The generated images are evaluated using multiple quality assessment techniques. Structural consistency is examined using contour-based analysis, while color similarity is assessed through histogram comparison. Texture similarity is measured using template matching approaches, and noise levels are quantified using Laplacian-based analysis. These evaluation metrics guide model optimization and ensure the generation of photorealistic facial images with reduced artifacts and improved clarity.

Implementation

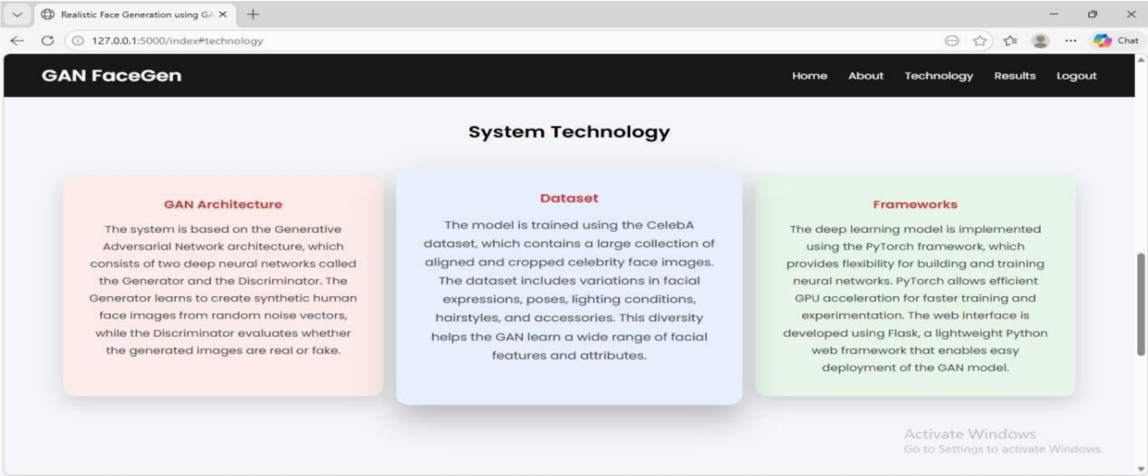
The implementation of the proposed system utilizes several Python-based libraries to support deep learning development and deployment. PyTorch serves as the primary deep learning framework, enabling construction of generator and discriminator networks, dynamic computation graphs, and efficient tensor operations. It also provides built-in loss functions and optimizers required for adversarial training. Torchvision complements PyTorch by offering image transformation utilities and dataset handling features, including resizing, normalization, and batch loading of the CelebA dataset.

NumPy is used for numerical computations and efficient manipulation of multi-dimensional arrays during preprocessing and evaluation. Pillow (PIL) handles image conversion and manipulation, allowing generated tensor outputs to be transformed into displayable image formats. Matplotlib is employed for visualizing training progress, loss curves, and generated images. Flask is used as a lightweight backend framework to deploy the trained model and handle user requests. Flask-Ngrok enables temporary public access to the locally hosted application for demonstration purposes. Gradio is utilized to create an interactive user interface, allowing users to generate facial images by selecting attributes such as age, smile, hair color, and gender. These libraries collectively support model training, visualization, and deployment.During training, random noise vectors are passed to the generator to create synthetic images. These generated samples, along with real images, are provided to the discriminator for classification. The generator and discriminator losses are calculated and backpropagated to update model weights iteratively. After training, the generator model is saved for future use. An interactive interface is implemented using Gradio, allowing users to modify facial attributes such as age, smile, hair color, and gender. The interface generates customized facial images in real time, demonstrating the effectiveness of the trained DCGAN model.

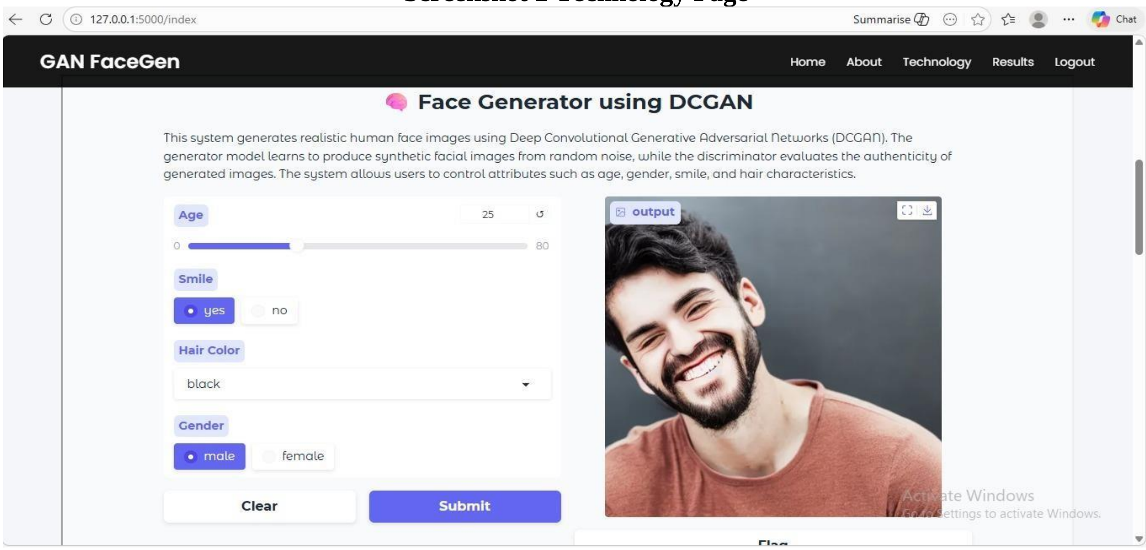
Screen Shots

Screenshot 1 Home Page

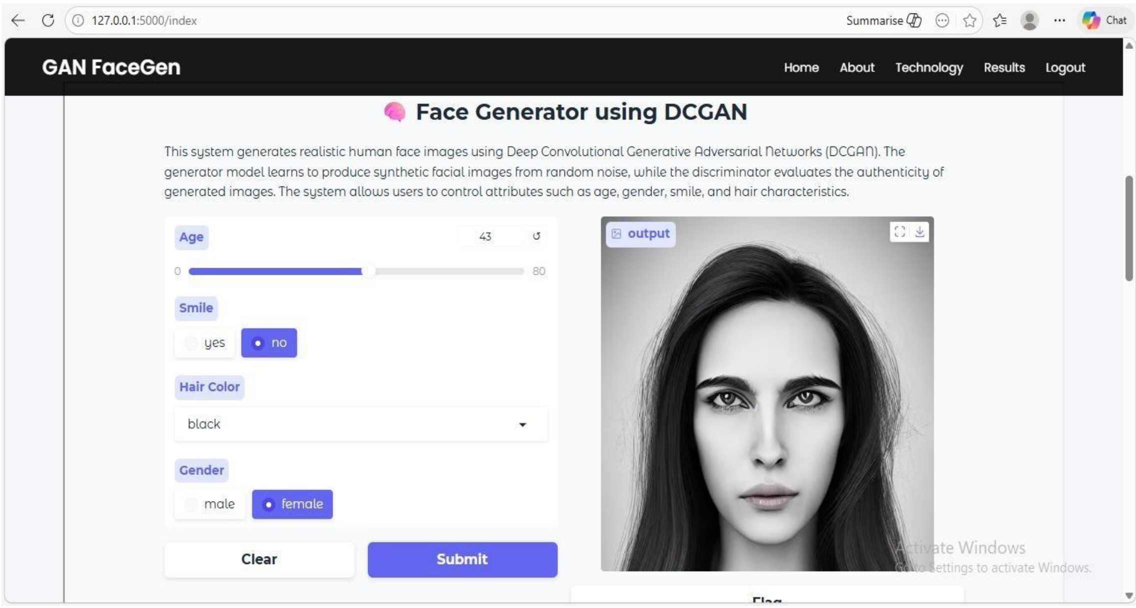
ISSN 2347–3657
Volume 14, Issue 1, 2026



Screenshot 2 Technology Page

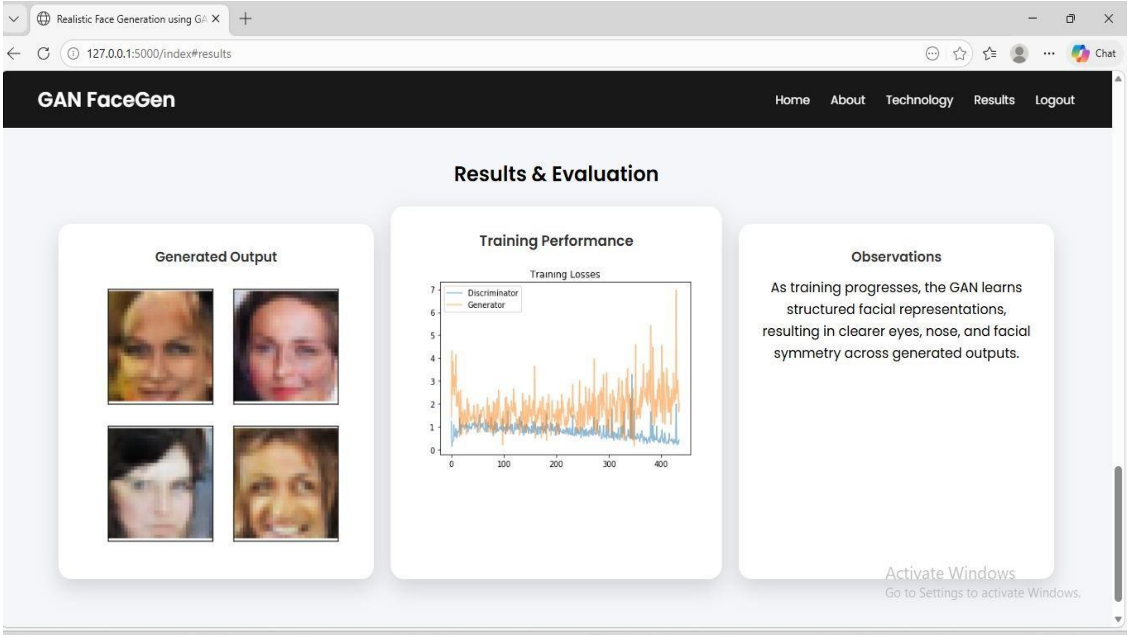


Screenshot 3 Results Page 1



Screenshot 4 Result Page 2

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 14, Issue 1, 2026



Screenshot 5 Result Page

Conclusion

This study presented a system for realistic human face generation using a Deep Convolutional Generative Adversarial Network (DCGAN), demonstrating the effectiveness of adversarial learning in synthesizing human-like images. The implemented architecture consists of generator and discriminator networks constructed with convolutional and transposed convolutional layers, enabling the model to learn complex facial features from a large-scale dataset. Through iterative adversarial training, the generator progressively improves its ability to produce visually convincing facial images, while the discriminator enhances its capacity to distinguish authentic images from synthetic outputs.To improve user interaction, the system incorporates controllable attributes such as age, smile, hair color, and gender, allowing customized facial image generation. During development, challenges including training instability, noise in generated images, and hardware limitations were encountered and addressed through appropriate preprocessing, normalization, and hyperparameter tuning. The use of PyTorch facilitated efficient model development, while frameworks for interface integration enabled smooth interaction between backend processing and user input. Although the generated images may lack perfect detail due to limited computational resources and training duration, the results demonstrate that DCGAN-based models can produce realistic facial images suitable for practical applications.The study also highlights the importance of proper dataset preparation, parameter optimization, and architectural design in improving model performance. The proposed system provides a foundation for further research in generative modeling and interactive image synthesis. Potential applications include entertainment systems, virtual avatar creation, gaming environments, and dataset augmentation for machine learning tasks. Ethical considerations related to the misuse of synthetic faces are acknowledged, emphasizing the need for responsible deployment and monitoring in future developments.

Future Scope

The proposed DCGAN-based face generation system offers several opportunities for future enhancements. One possible extension is the implementation of conditional generative adversarial networks to enable more precise control over facial attributes such as age, emotion, hairstyle, and accessories. Real-time face editing capabilities can also be incorporated, allowing users to dynamically modify features without regenerating images from scratch. Another promising direction involves face-to-face translation, where an input image can be transformed into multiple variations with different attributes.Further improvements may include the development of three-dimensional face generation models, which would be beneficial for virtual reality, gaming, and animation applications. The system can also be deployed as a full-scale web or mobile application to improve accessibility and user experience. Cloud-based deployment can enhance scalability, reliability, and continuous availability. Additionally, generated images may be utilized for data augmentation to improve the performance of other machine learning models.Future work may also integrate face recognition or verification modules to create a comprehensive facial analysis platform. Enhancements in training strategies and architectural design can reduce noise and improve stability during adversarial learning. Ethical safeguards, such as incorporating deepfake detection mechanisms and responsible usage policies, should be implemented to prevent misuse of generated facial images. Overall, the proposed system provides a strong foundation for developing more advanced and practical artificial intelligence applications in image synthesis and computer vision.

References

[1] M. Manimegala, V. Gokulraj, K. Karisni, and S. Manisha, "Generating Human Face with DCGAN and GAN," International Research Journal on Advanced Engineering Hub, vol. 2, no. 5, pp. 1348–1354, May 2024.

[2] A. Garg, "Parameter and Data-Efficient Spectral StyleDCGAN," arXiv preprint arXiv:2404.00597, 2024.

[3] N. Shukla and A. Ross, "dc-GAN: Dual-Conditioned GAN for Face Demorphing From a Single Morph," arXiv preprint arXiv:2411.14494, 2024.

[4] J. Lu, "Enhanced Anime Image Generation Using USE-CMHSA-GAN," arXiv preprint arXiv:2411.11179, 2024.

[5] X. Wang, H. Guo, S. Hu, M.-C. Chang, and S. Lyu, "GAN-generated Faces Detection: A Survey and New Perspectives," IEEE Access / Survey, 2023.

[6] "Application of Generative Adversarial Networks in Image and Face Generation," IEEE- related Journal / Taylor & Francis, 2024.

[7] "Exploring Deep Convolutional Generative Adversarial Networks for Biometric Systems," Springer / IEEE-related research, 2024.