IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

SYNTHETIC HUMAN TWIN: AN AI - POWERED BEHAVIORAL REPLICATION SYSTEM

M. Asha[1], A. Sri Malleswari[2], B. Prema Chandana[3], D. Tabitha[4]

Assistant Professor, Department of Computer Science and Engineering (AIML),

Bapatla Women's Engineering College, Bapatla – 522101, Andhra Pradesh, India[1]

Students, Department of Computer Science and Engineering (AIML),

Bapatla Women's Engineering College, Bapatla – 522101, Andhra Pradesh, India[2-4]

Abstract: Facial expression manipulation using artificial intelligence has gained significant attention in recent years due to its applications in digital media, content creation, and virtual identity systems. However, existing AI-based image editing techniques often suffer from the problem of *identity drift*, where the generated output alters key facial features such as skin tone, eye shape, or facial structure, resulting in images that no longer resemble the original subject. This paper introduces Image Replica (Expression AI), a prompt-driven AI system designed to transform facial expressions in portrait images while preserving the subject's identity.

The proposed system integrates a modern full-stack web architecture consisting of a Next.js frontend and a Fast API backend, combined with advanced generative AI models. A novel **Identity-Preserving Prompt Engine** is implemented to enrich user prompts with structural and photorealistic constraints before being processed by diffusion-based image synthesis models. The system utilizes Google's Gemini model for prompt interpretation and OpenAI's DALL-E model for high-quality image generation.

To ensure privacy, the platform adopts an in-memory processing pipeline that prevents any persistent storage of user images. Experimental results demonstrate that the proposed approach produces highly photorealistic expression transformations while maintaining strong identity consistency across generated outputs. The system provides a scalable, accessible, and privacy-conscious solution for AI-driven facial expression synthesis.

Keywords: Generative Artificial Intelligence, Facial Expression Synthesis, Identity Preservation, Prompt Engineering, Diffusion Models, Image-to-Image Transformation, Privacy-Preserving AI, Full-Stack AI Architecture, Digital Twin Generation.

I. INTRODUCTION

Facial expressions play a vital role in human communication by conveying emotions, intentions, and social cues. In the era of digital media, visual identity and expressive imagery have become increasingly important in domains such as social media, digital marketing, and virtual identity platforms. Consequently, the ability to modify facial expressions in images while preserving the original identity of the subject has become an important capability for modern digital content creation.

Traditional image editing tools such as Adobe Photoshop and GIMP require significant manual effort and professional expertise to achieve realistic facial expression modifications. Although several mobile applications and AI-based filters attempt to simplify this process, they often rely on predefined transformations that produce unrealistic results. A major challenge in many AI-based image editing systems is **identity drift**, where the generated image alters key facial features such as facial structure, skin tone, or lighting conditions, causing the output to deviate from the original subject.

Recent advancements in **Generative Artificial Intelligence**, particularly diffusion models and multimodal vision–language systems, have significantly improved the ability to generate photorealistic images from textual descriptions. However, ensuring precise expression editing while maintaining identity consistency remains a challenging problem.
To address this issue, this paper proposes **Image Replica (Expression AI)**, a prompt-driven AI system that enables identity-preserving facial expression transformation. The platform integrates a Next.js frontend, a FastAPI backend, and generative AI models such as Google's Gemini and OpenAI's DALL-E to generate photorealistic results. Additionally,

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

a privacy-first in-memory processing pipeline ensures that user images are not stored, improving data security and user privacy.

II. LITERATURE REVIEW AND PROBLEM IDENTIFICATION

Facial expression synthesis and manipulation have been widely studied in the fields of computer vision and generative artificial intelligence. Over the past decade, researchers have explored various techniques for modifying facial attributes in images, including traditional image editing tools, filter-based applications, and more recently, deep learning-based generative models.

Early approaches to facial editing relied heavily on manual image manipulation using professional software such as Adobe Photoshop and GIMP. These tools provide precise control over image elements but require significant expertise and time to achieve realistic results. To simplify the process, several mobile and web-based applications introduced filter-based techniques that automatically apply predefined transformations to facial images. Applications such as FaceApp and Snapchat filters allow users to modify facial expressions, apply aging effects, or add stylistic transformations. However, these systems primarily rely on fixed transformation templates and lack semantic understanding of user intent, often producing unnatural or overly stylized results.

With the emergence of deep learning, **Generative Adversarial Networks (GANs)** significantly advanced the field of facial image synthesis. Introduced by Goodfellow et al. (2014), GAN-based models enabled the generation of realistic human faces and attribute manipulation. Later research introduced architectures such as StarGAN and StyleGAN, which allowed conditional image generation and multi-domain facial attribute editing. Although GAN-based systems produce impressive results, they often suffer from training instability, mode collapse, and limited controllability when attempting fine-grained expression editing.

More recently, **diffusion-based generative models** have emerged as a powerful alternative to GANs for high-quality image generation. Diffusion models, such as those used in OpenAI's DALL-E and Stable Diffusion systems, generate images by iteratively denoising random noise guided by textual prompts. These models demonstrate superior photorealism and greater flexibility in text-guided image synthesis. Additionally, multimodal vision-language models such as Google's Gemini have improved the ability of AI systems to interpret natural language instructions and visual context simultaneously.

Despite these advancements, existing systems still face several critical challenges when applied to facial expression editing. One of the most significant issues is **identity drift**, where the generative model unintentionally alters core facial characteristics such as facial structure, eye color, skin tone, or lighting conditions. As a result, the generated image may no longer resemble the original subject, limiting the practical usefulness of the transformation.

Another limitation is the lack of precise **prompt-controlled editing mechanisms** that ensure only specific facial components related to expression are modified. Many generative systems treat the task as a full image synthesis problem rather than controlled modification of an existing photograph. Furthermore, privacy concerns arise because many AI platforms store user-uploaded images on remote servers for processing, creating potential risks when handling biometric data such as human faces.

Therefore, there is a need for an intelligent system capable of performing **identity-preserving facial expression synthesis** while maintaining photorealistic quality and strong user control through natural language prompts. This work addresses these challenges by proposing **Image Replica (Expression AI)**, a prompt-driven facial expression transformation system that integrates identity-preserving prompt engineering, modern diffusion-based image generation models, and a privacy-first processing architecture.

III. SYSTEM DESIGN AND METHODOLOGY

The proposed system, **Image Replica (Expression AI)**, is designed to transform facial expressions in portrait images while preserving the subject's identity. The system follows a modular and scalable architecture that integrates modern web technologies with advanced generative AI models. The overall methodology focuses on prompt-driven expression synthesis, identity preservation, and privacy-conscious data handling.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

A. System Architecture

The platform follows a **decoupled client–server architecture** consisting of three primary layers: the presentation layer, the processing layer, and the AI model layer.

The **presentation layer** is implemented using Next.js and React, which provide a responsive web interface that allows users to upload portrait images and specify the desired facial expression through natural language prompts or preset options. This layer is responsible for user interaction, image preview generation, prompt input handling, and displaying the generated results.

The **processing layer** is implemented using the FastAPI framework running on a Uvicorn ASGI server. This backend layer receives the uploaded image and expression prompt from the frontend via multipart HTTP requests. The backend performs input validation, prompt enrichment, and communication with the external AI image generation service.

The **AI model layer** consists of cloud-based generative models responsible for image synthesis. Google's Gemini model is used for interpreting and refining the user prompt, while OpenAI's DALL-E model performs the final high-resolution image generation. This combination enables both semantic understanding of user instructions and photorealistic image synthesis.

B. Image Processing Workflow

The system processes each request through a structured workflow. First, the user uploads a portrait image through the web interface. The frontend validates the file format and size to ensure compatibility with the system requirements. After validation, the user enters a textual prompt describing the desired facial expression.

The frontend then constructs a multipart request containing the image file and the prompt, which is sent to the FastAPI backend. Upon receiving the request, the backend reads the image as a byte stream and passes it to the prompt enrichment module.

Next, the enriched prompt and image data are forwarded to the AI image generation model. The generative model processes the input and produces a new image with the modified facial expression. The generated image is returned as a binary stream and rendered in the user interface, allowing the user to view and download the result.

C. Identity-Preserving Prompt Engineering`

A major challenge in AI-based facial editing systems is maintaining the identity of the original subject during expression modification. To address this challenge, the proposed system implements an **Identity-Preserving Prompt Engine**.

The prompt engine operates using a two-layer prompt composition strategy. The first layer is a **system-defined base prompt** that instructs the generative model to maintain photorealism, preserve facial structure, and ensure consistent lighting conditions. The second layer is the **user-defined prompt**, which specifies the desired facial expression.

These two components are combined to form the final prompt that is sent to the image generation model. By enforcing structural constraints within the base prompt, the system ensures that only the expressive regions of the face are modified while preserving the subject's identity.

D. Privacy-Preserving Data Processing

Since the system processes biometric data in the form of human face images, privacy protection is a key design requirement. The proposed platform adopts a **transient in-memory processing model** to eliminate the need for persistent data storage.

When a user uploads an image, the backend processes the file directly in memory using byte-stream buffers. The image is never written to disk or stored in a database. After the AI generation process is completed and the result is delivered to the user, the image data is immediately discarded from memory. This design ensures that no personal image data is retained on the server, providing a privacy-first AI processing pipeline.

Through this system design and methodology, the proposed platform enables efficient, scalable, and identity-consistent facial expression synthesis while maintaining strong privacy guarantees.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

Figure 1: Architecture of the AI-Driven Sugarcane Disease Detection and Recommendation System

IV. PEER REVIEW AND VALIDATION

To ensure the reliability and effectiveness of the proposed Image Replica (Expression AI) system, the platform was evaluated through peer review and validation procedures. The evaluation focused on assessing the system's performance in terms of photorealism, identity preservation, usability, and overall system reliability.

The system was tested by a group of peer reviewers consisting of students and researchers with knowledge in artificial intelligence and web-based systems. Participants were asked to upload portrait images, apply different facial expression prompts, and analyze the generated results. The reviewers evaluated whether the generated images maintained the identity of the original subject while accurately reflecting the requested facial expressions.

The validation process considered several evaluation criteria, including the realism of generated images, consistency of facial features, system responsiveness, and ease of use of the interface. The majority of reviewers reported that the generated outputs closely resembled the original subject and successfully reflected the desired expressions. The identity-preserving prompt strategy effectively minimized identity drift while maintaining natural lighting and facial structure.
In addition, usability testing indicated that the web-based interface was intuitive and easy to operate. The prompt input mechanism and preset expression options simplified user interaction, allowing users to generate results without requiring technical expertise.

Overall, the peer review results confirmed that the proposed system performs reliably and produces photorealistic expression transformations while maintaining strong identity consistency. The evaluation demonstrates that Image Replica (Expression AI) is a practical and scalable solution for AI-driven facial expression synthesis.

V. RESULTS AND DISCUSSION

The implementation of the **Image Replica (Expression AI)** platform demonstrates the successful integration of generative artificial intelligence with modern web-based system architecture. The system was evaluated based on visual realism, identity preservation, system responsiveness, and usability.

The first stage of the system is the landing interface where users can access the platform and begin the image transformation process. As shown in **Figure 1**, the landing page provides a clean and user-friendly interface that introduces the system functionality and allows users to start uploading images for expression transformation.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

Figure 1: Landing Page of the Image Replica System

The next stage involves the image upload interface. Users can upload a portrait image and enter a textual description of the desired facial expression. **Figure 2** illustrates the uploading page where the user selects an image and provides the expression prompt that will guide the AI generation process.

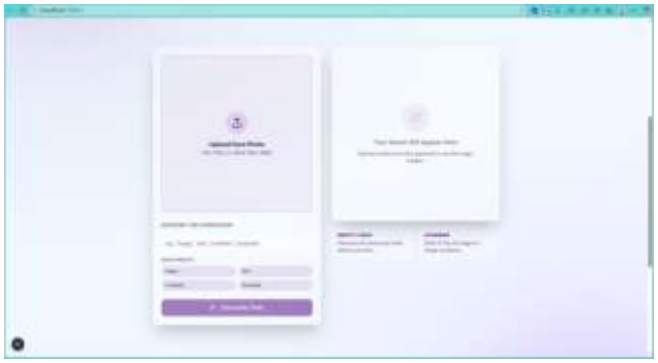

Figure 2: Image Uploading Interface

Once the image is successfully uploaded, the system displays the input image and allows the user to specify the expression to be generated. As shown in **Figure 3**, the uploaded image is previewed within the interface, and the user can select preset expressions or enter custom prompts.

Figure 3: Input Image Uploaded with Expression Selection

After the generation process is completed, the AI model produces a photorealistic output image that reflects the requested facial expression while maintaining the identity of the original subject. **Figure 4** demonstrates the final output generated by the system. The results show that the system successfully modifies facial expressions while preserving facial structure, skin tone, and lighting conditions.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2026

DOI: 10.17148/IARJSET.2026.134127

Figure 4: AI Generated Output Image

The experimental results indicate that the proposed system produces visually realistic images with minimal identity drift. The prompt-driven architecture enables flexible control over facial expressions while maintaining photorealism. In addition, the privacy-first processing pipeline ensures that user images are handled securely without persistent storage. Overall, the results confirm that the **Image Replica (Expression AI)** platform effectively performs identity-preserving facial expression transformation while providing an intuitive user experience and reliable system performance.

VI. CONCLUSION

This study presented **Image Replica (Expression AI)**, an AI-powered system designed to perform facial expression transformation while preserving the identity of the original subject. The proposed platform integrates modern web technologies with advanced generative AI models to create a scalable and user-friendly solution for image-based emotion synthesis.

The system combines a **Next.js frontend**, **FastAPI backend**, and cloud-based generative models such as **OpenAI's DALL-E** and **Google's Gemini** to enable prompt-driven facial expression editing. A key component of the system is the **Identity-Preserving Prompt Engine**, which enriches user prompts with structural constraints to ensure that the generated images maintain the original subject's facial structure, lighting conditions, and visual characteristics. This approach effectively addresses the common problem of **identity drift** observed in many existing AI image editing systems.

Experimental results demonstrate that the proposed platform is capable of producing **photorealistic facial expression transformations** while maintaining identity consistency. In addition, the system's **privacy-first in-memory processing architecture** ensures that user images are processed securely without being permanently stored on the server.

Overall, the Image Replica platform provides an efficient and accessible solution for AI-driven facial expression synthesis. The integration of prompt engineering, generative diffusion models, and modern full-stack architecture highlights the potential of combining artificial intelligence with web technologies to develop advanced digital content creation tools.

Future Enhancements

Although the proposed **Image Replica (Expression AI)** platform successfully demonstrates identity-preserving facial expression transformation, several improvements can further enhance its capabilities and expand its practical applications.

One potential enhancement is the integration of **true image-to-image editing techniques** using mask-based editing models. This would allow the system to directly modify specific facial regions of the uploaded image rather than generating a new image from prompts, thereby improving identity accuracy and producing even closer replicas of the original subject.

Another important improvement is the addition of **real-time webcam integration**, which would enable users to capture images directly within the platform. This feature would simplify the user workflow and provide a more seamless experience for generating expression transformations.The system can also be extended to support **multilingual prompt input**, allowing users to describe facial expressions in different languages. Integrating a language translation layer would make the platform more accessible to users from diverse linguistic backgrounds.

DOI: 10.17148/IARJSET.2026.134127

Furthermore, future work may explore **video-based facial expression transformation**, where expressions can be modified frame-by-frame in short video clips. This capability would open new possibilities in digital content creation, film production, and virtual communication.

Finally, integrating **direct social media publishing features** could allow users to share generated images instantly on platforms such as Instagram, LinkedIn, or Twitter. These enhancements would transform the Image Replica platform into a more comprehensive AI-powered creative tool for digital identity and visual content generation.

ACKNOWLEDGEMENT

The successful completion of this project would not have been possible without the support and guidance of several individuals and institutions. The authors would like to express their sincere gratitude to the faculty members and mentors who provided valuable insights, technical guidance, and continuous encouragement throughout the development of this work.

We would also like to thank our department and institution for providing the necessary resources, infrastructure, and learning environment that enabled us to carry out this research successfully.

Special appreciation is extended to the developers and researchers behind the open-source frameworks and technologies used in this project, including **Next.js, React, FastAPI, Tailwind CSS**, and the AI models provided by **OpenAI and Google**. These tools played a significant role in enabling the development and implementation of the Image Replica (Expression AI) platform.

Finally, we would like to acknowledge the peers and participants who contributed their time and feedback during the testing and validation phase, helping us improve the functionality and usability of the system.

Their support and encouragement have been instrumental in the successful completion of this work.

REFERENCES

[1]. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Networks," *Advances in Neural Information Processing Systems (NeurIPS)*, 2014.
[2]. J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
[3]. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
[4]. OpenAI, "DALL·E: Creating Images from Text," OpenAI Research, 2023. Available: https://openai.com/research/dall-e
[5]. Google AI, "Gemini: A Family of Highly Capable Multimodal Models," Google DeepMind Research, 2024.
[6]. FastAPI Documentation. Available: https://fastapi.tiangolo.com/
[7]. Next.js Documentation. Available: https://nextjs.org/docs
[8]. React Documentation. Available: https://react.dev/
[9]. Tailwind CSS Documentation. Available: https://tailwindcss.com/docs
[10]. A. Radford et al., "Learning Transferable Visual Models from Natural Language Supervision," *Proceedings of the International Conference on Machine Learning (ICML)*, 2021.
[11]. T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
[12]. M. Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning," *USENIX Symposium on Operating Systems Design and Implementation*, 2016.

