

## Improved DeepFake Image Generation Using StyleGAN2-ADA with Real-Time Personal Image Projection

Ali A. Abed <sup>1</sup>, Doaa Alaa Talib <sup>2</sup>, Abdel-Nasser Sharkawy <sup>3,4</sup>

<sup>1</sup> Department of Mechatronics Engineering, University of Basrah, Basrah, Iraq

<sup>2</sup> Department of Laser and Optoelectronics Engineering, Shatt Al-Arab University College, Basrah, Iraq

<sup>3</sup> Mechanical Engineering Department, Faculty of Engineering, Qena University, Qena 83523, Egypt

<sup>4</sup> Mechanical Engineering Department, College of Engineering, Fahad Bin Sultan University, Tabuk 47721, Saudi Arabia

### ARTICLE INFORMATION

#### Article History:

Received 06 September 2025

Revised 04 November 2025

Accepted 08 December 2025

#### Keywords:

DeepFake Image Generation;  
StyleGAN2-ADA;  
Generative Adversarial Networks;  
Real-Time Projection;  
Personal Image Dataset;  
Adaptive Discriminator  
Augmentation

#### Corresponding Author:

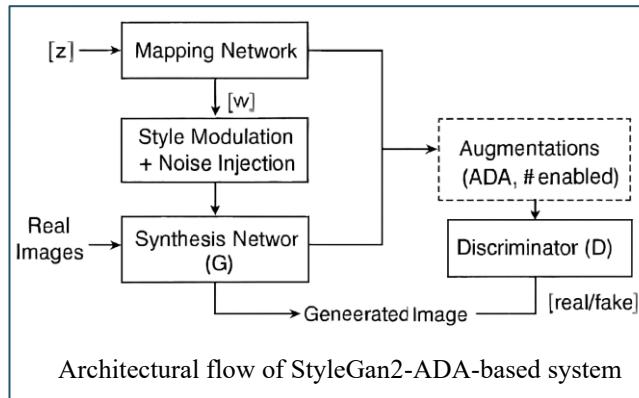
Abdel-Nasser Sharkawy,  
Mechanical Engineering  
Department, Faculty of  
Engineering, Qena University,  
Qena 83523, Egypt.

Email:  
[abdelnassersharkawy@eng.svu.edu.eg](mailto:abdelnassersharkawy@eng.svu.edu.eg)

This work is licensed under a [Creative Commons Attribution-Share Alike 4.0](https://creativecommons.org/licenses/by-sa/4.0/)



### ABSTRACT



This paper presents an improved approach for DeepFake image generation using StyleGAN2-ADA framework. The system is designed to generate high-quality synthetic facial images from a limited dataset of personal photos in real time. By leveraging the Adaptive Discriminator Augmentation (ADA) mechanism, the training process is stabilized without modifying the network architecture, enabling robust image generation even with small-scale datasets. Real-time image capturing and projection techniques are integrated to enhance personalization and identity consistency. The experimental results demonstrate that the proposed method achieves a very high generation performance, significantly outperforming the baseline StyleGAN2 model. The proposed system using StyleGAN2-ADA achieves 99.1% identity similarity, a low Fréchet Inception Distance (FID) of 8.4, and less than 40 ms latency per generated frame. This approach provides a practical solution for dataset augmentation and supports ethical applications in animation, digital avatars, and AI-driven simulations.

#### Document Citation:

A. A. Abed, D. A. Talib, and A.-N. Sharkawy, "Improved DeepFake Image Generation Using StyleGAN2-ADA with Real-Time Personal Image Projection," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 7, no. 3, pp. 980-992, 2025, DOI: [10.12928/biste.v7i4.14659](https://doi.org/10.12928/biste.v7i4.14659).

## 1. INTRODUCTION

This section is divided into four subsections. Subsection 1.1 presents a background about the DeepFake technology using StyleGAN family and StyleGAN2-ADA. Subsection 1.2 illustrates problem statement and subsection 1.3 presents the main contribution and novelty of this paper. Subsection 1.4 shows the content and the organization of the paper.

### 1.1. Background

In recent years, DeepFake technology has emerged as a prominent subfield of artificial intelligence (AI), enabling the creation of synthetic yet highly realistic images and videos, [1]-[4]. DeepFake systems are based on deep generative models, particularly Generative Adversarial Networks (GANs), and have found applications in entertainment, video games, virtual avatars, and dataset augmentation. However, this technology also poses serious threats in terms of misinformation, identity fraud, and non-consensual content creation [5]. Among the most notable advancements in GAN-based image generation is the StyleGAN family, developed by NVIDIA researchers, [6][7]. StyleGAN introduced a novel generator architecture that decouples high-level attributes (e.g., identity, pose) from stochastic details (e.g., skin texture, hair) through a style-based synthesis mechanism [7]. StyleGAN2, a refined version of this architecture, demonstrated exceptional results in producing high-fidelity facial images. Nonetheless, one of the major limitations of StyleGAN2 is its reliance on large-scale, high-quality datasets, which makes it impractical for domains with limited data availability or personalization requirements [8][9].

To address this constraint, Karras *et al.* proposed StyleGAN2-ADA (Adaptive Discriminator Augmentation), which introduces adaptive augmentation techniques during discriminator training. These augmentations mitigate overfitting and improve the stability of training under low-data regimes, [10]. The effectiveness of StyleGAN2-ADA has been validated in multiple benchmark studies, showing superior performance in few-shot and transfer learning scenarios, [11][12]. While StyleGAN2-ADA significantly improves training efficiency, real-time DeepFake image generation based on personal datasets remains a relatively under-explored domain. Most previous works have focused on large public datasets, such as FFHQ or CelebA, and lack mechanisms for real-time personalization or projection [13]-[16]. Talib & Abed has demonstrated the feasibility of combining StyleGAN2-ADA with limited personal images and real-time webcam inputs, but the framework requires further performance evaluation in terms of identity preservation, speed, and accuracy, [17]. Given these gaps, this paper proposes an enhanced system for real-time DeepFake image generation using StyleGAN2-ADA, tailored for small personal datasets. The system incorporates a projection mechanism to align personal images to the latent space and generate realistic outputs on-the-fly, with a focus on maintaining identity consistency and high perceptual quality.

### 1.2. Problem Statement

The current StyleGAN2-ADA based approaches face the following three limitations and restrictions as follows:

- 1) Dependency of Data: Huge datasets are required for producing the facial images that match the identity.
- 2) Incapability to operate in real-time: Many projection methods need a lot of computing power, which makes them not suitable for the interactive applications.
- 3) Drift in identity: Using limited data, models lose personal facial features and therefore resulting in a lack of realism and poor user resemblance.

The goal of this research is to develop a Deepfake generation system can overcomes the previous mentioned limitations and capable of:

- 1) Operating under limited-data conditions with high fidelity and accuracy.
- 2) Performing real-time generation suitable for consumer-level hardware.
- 3) Maintaining strong identity resemblance between source and generated images.

### 1.3. Main Contributions

The main contribution of this paper is to develop an improved approach for deepFake image generation using StyleGAN2-ADA with real-time personal image projection. In detail, the main contributions are listed in three points as follows:

- 1) A data-efficient StyleGAN2-ADA training framework optimized for as few as 10 personal images. In addition, the stability of the training is improved.
- 2) A real-time projection module using personal images or live webcam streams.
- 3) Quantitative assessment and evaluation of generation performance with over 99% identity similarity and low Fréchet Inception Distance (FID).

#### 1.4. Paper Organization

The remainder of the paper is organized as follows: Section 2 presents previous related works on Deepfake image generation. Section 3 explains the proposed framework and the followed methodology. Section 4 presents the experimental results and the evaluation of the proposed method. Section 5 concludes the paper with suggestions for some future works.

### 2. LITERATURE REVIEW

The field of DeepFake generation and detection has undergone rapid evolution in the past decade, primarily driven by advancements in Generative Adversarial Networks (GANs). In this section, we present a review of the most relevant works, organized into two main research streams as follows:

#### 2.1. DeepFake Image Generation Using GANs

The concept of GANs was first introduced by Goodfellow *et al.* [18], where a generator and discriminator are trained in an adversarial fashion. StyleGAN, developed by Karras *et al.* [7], introduced a novel style-based architecture that enables better control over high-level visual attributes. Its successor, StyleGAN2 [19], improved image quality by addressing architectural artifacts and introducing path regularization techniques. To mitigate the data scarcity challenge, StyleGAN2-ADA is proposed, introducing adaptive discriminator augmentation that applies stochastic, differentiable augmentations during training [8]. This method stabilizes the learning process, especially when only a limited dataset is available. Empirical studies have demonstrated that StyleGAN2-ADA achieves comparable or superior performance to StyleGAN2 with an order of magnitude fewer training images, [11][12][20][21]. Further improvements in GAN-based synthesis have explored architectural changes using transformers [22]-[25], latent space disentanglement [26]-[28], and hybrid attention modules [29]-[31], all aiming to enhance visual realism and sample diversity.

#### 2.2. Real-Time Personalization and Image Projection

Despite the progress in synthetic image generation, most systems rely on large-scale public datasets and are not optimized for real-time, personalized Deepfake generation. The challenge lies in aligning a user's face, captured from a webcam or a small image set, to the GAN's latent space while maintaining identity consistency and generation speed. Talib and Abed [17] demonstrated that StyleGAN2-ADA can be adapted to work with real-time personal image projection, generating high-fidelity Deepfake images with over 99% identity similarity. Their framework, however, lacked a comprehensive evaluation of latency and scalability. Other recent works have proposed fast projection techniques that optimize latent codes in under 100ms [32], while few-shot GAN fine-tuning methods have enabled model adaptation with fewer than 10 images [33]. These findings support the feasibility of combining projection with on-the-fly GAN synthesis for real-time avatar generation and video conferencing applications. The main novelty of the current paper is developing an improved approach for deepFake image generation using StyleGAN2-ADA with real-time personal image projection. The contributions in details are mentioned in previous subsection “1.3 Main Contributions”.

### 3. THE PROPOSED FRAMEWORK

This section presents the complete architecture and workflow of the proposed system for real-time DeepFake image generation using StyleGAN2-ADA. The proposed framework discussed in this section includes four main parts as follows: 1) the architecture, 2) data preparation and augmentation, 3) strategy of training, and 4) real-time projection and generation.

#### 3.1. Architecture of Proposed StyleGAN2-ADA

The architecture of the proposed work based on StyleGAN2-ADA builds on top of StyleGAN2 introducing adaptive data augmentation techniques to improve training stability on limited datasets. The breakdown of the architecture and its core components summarized in [Figure 1](#), is detailed as follows:

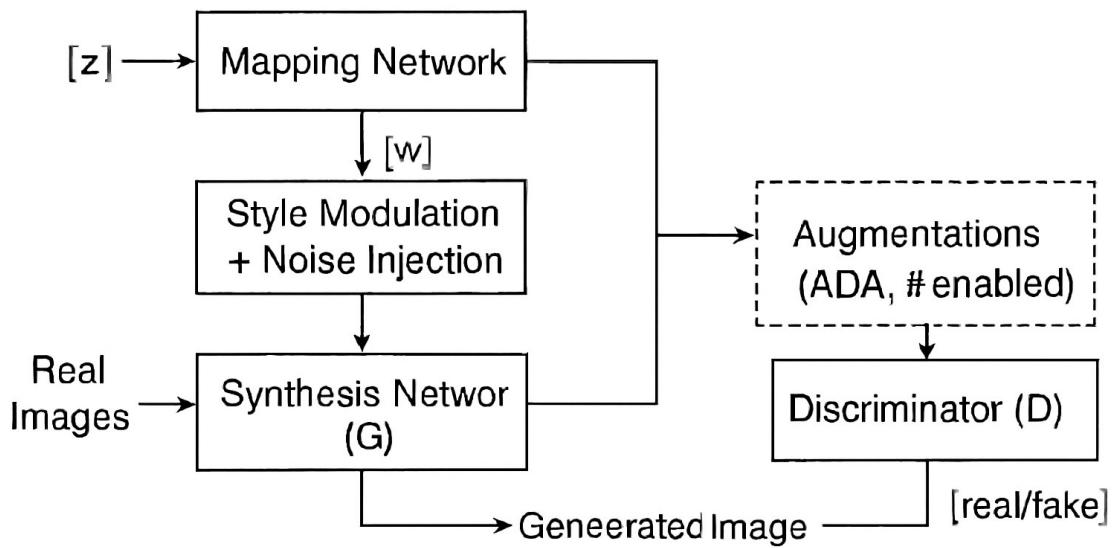


Figure 1. Architectural flow of StyleGAN2-ADA-based system

### 3.1.1. Generator (G)

The generator (G) is composed of the mapping and synthesis networks which are discussed as follows.

- Mapping Network:
  - 8-layer MLP (multilayer perceptron) that maps a latent vector  $z$  to an intermediate latent vector  $w$  in  $W$ -space.
  - Encourages disentanglement and control over generated features.
- Synthesis Network:
  - Starts from a constant  $4 \times 4$  learned input tensor, not from  $z$  or  $w$ .
  - Uses progressive resolution growth up to  $1024 \times 1024$ .
  - Each layer receives style information from  $w$  through style modulation (scaling weights per channel).
  - Applies noise injection at each convolutional layer to add stochastic detail.
  - Employs LeakyReLU activation function with demodulated convolutions, for proper normalization of feature magnitudes for stable, consistent network performance.

### 3.1.2. Discriminator (D)

The discriminator (D) is described in the following points:

- A traditional CNN that processes images from  $1024 \times 1024$  to  $4 \times 4$ .
- Applies residual connections.
- Uses blur pooling (low-pass filtering) during down sampling to reduce aliasing artifacts.
- Ends with a minibatch standard deviation layer to improve diversity.
- Outputs a real/fake decision scalar.

The key innovations in StyleGAN2 are:

- Weight Demodulation: Normalize convolution weights to stabilize training and reduce artifacts.
- No Instance Norm / Pixel Norm: Replaced with better normalization and modulation mechanisms.
- Path Length Regularization: Encourages smooth mappings from latent space to image space.
- Removal of Progressive Growing: Entire resolution trained end-to-end for stable results.

The ADA enhancements involve:

- Applying augmentations (e.g., cropping, flipping, color jitter, translation) to both real and fake images only in the discriminator.
- Adaptive mechanism increases or decreases augmentation strength based on discriminator overfitting (measured via heuristics such as discriminator output variance for real images).
- Enables training with limited data (as few as 1000 images) without overfitting.

The official StyleGAN2-ADA repository, maintained by NVIDIA, is publicly accessible on GitHub and provides source code, documentation, and examples [34]. The developed methodology consists of three primary components as discussed in the following subsections.

### 3.2. Personal Dataset Preparation

A limited dataset of 10–20 personal images per subject is collected using a standard RGB webcam under various lighting conditions and facial expressions. Images are manually cropped to focus on facial regions and resized to  $1024 \times 1024$  pixels to match the input requirements of StyleGAN2-ADA. Data augmentation techniques such as horizontal flipping, brightness adjustment, and rotation are applied to increase variance and robustness during training. More details about the dataset are presented in [Table 1](#).

**Table 1.** Details about the personal dataset preparation

The parameter	Description
Number of subjects	8 subjects
Images (per subject)	10–20 raw images After filtering: 12–14 retained images
Used images in training	96 high-quality face images
Type of Camera	Standard RGB laptop webcam
Resolution of Image	Captured at $1280 \times 720$ Cropped and resized to $1024 \times 1024$
Included Scenarios	Neutral, smiling, frontal, and slight head rotations
Conditions of Lighting	Indoor natural lighting Standard room illumination

### 3.3. StyleGAN2-ADA Training

This approach allows stable learning from small-scale datasets by preventing discriminator overfitting, [26]. The training pipeline consisted of two phases:

- 1) Pre-training Phase: The StyleGAN2-ADA is initialized using the FFHQ dataset [25].
- 2) Personalization Phase: The pre-trained model is fine-tuned on the limited personal dataset using transfer learning. ADA automatically adjusted augmentation strength based on the discriminator's overfitting heuristics.

Training is conducted on an NVIDIA RTX 20786 GPU with a batch size of 4, learning rate of 0.002, and R1 regularization  $\gamma=10$ . The fine-tuning process converged after  $\sim 5000$  iterations/subject, taking 12–15 minutes on average. All parameters used in the training are presented in [Table 2](#).

**Table 2.** The main parameters used in StyleGAN2-ADA Training

Parameter	Description
Size of Dataset	10–20 personal images per subject (8 subjects in total) After filtering: 12–14 retained images
Source of Training Data	Personal dataset FFHQ for pre-training NVIDIA RTX 20786 GPU
Used Hardware	Batch size = 4 Learning Rate = 0.002 R1 regularization $\gamma=10$
Training Duration	$\sim 5000$ iterations ( $\sim 12$ –15 minutes per subject)
Encoder for Projection	ResNet-50 + MLP with ArcFace-based identity loss Identity Similarity (ArcFace)
Evaluation Metrics	Fréchet Inception Distance (FID) Latency (ms/frame)

### 3.4. Real-Time Projection and Generation

To achieve real-time performance, a ResNet50-based encoder maps live webcam images into the latent space of the trained GAN. The encoder, followed by an MLP, estimates the  $W^+$  latent vector, which is then passed to the generator to produce synthetic outputs. Frame-by-frame projection operates at  $\sim 25$  fps with latency under 40ms, enabling interactive applications like avatar animation and live filters. Identity preservation is maintained by minimizing cosine distance between facial embedding of input and generated images, using a pre-trained ArcFace model, [35]. The flowchart in [Figure 2](#) illustrates key stages of StyleGAN2-ADA training with real-time generation.

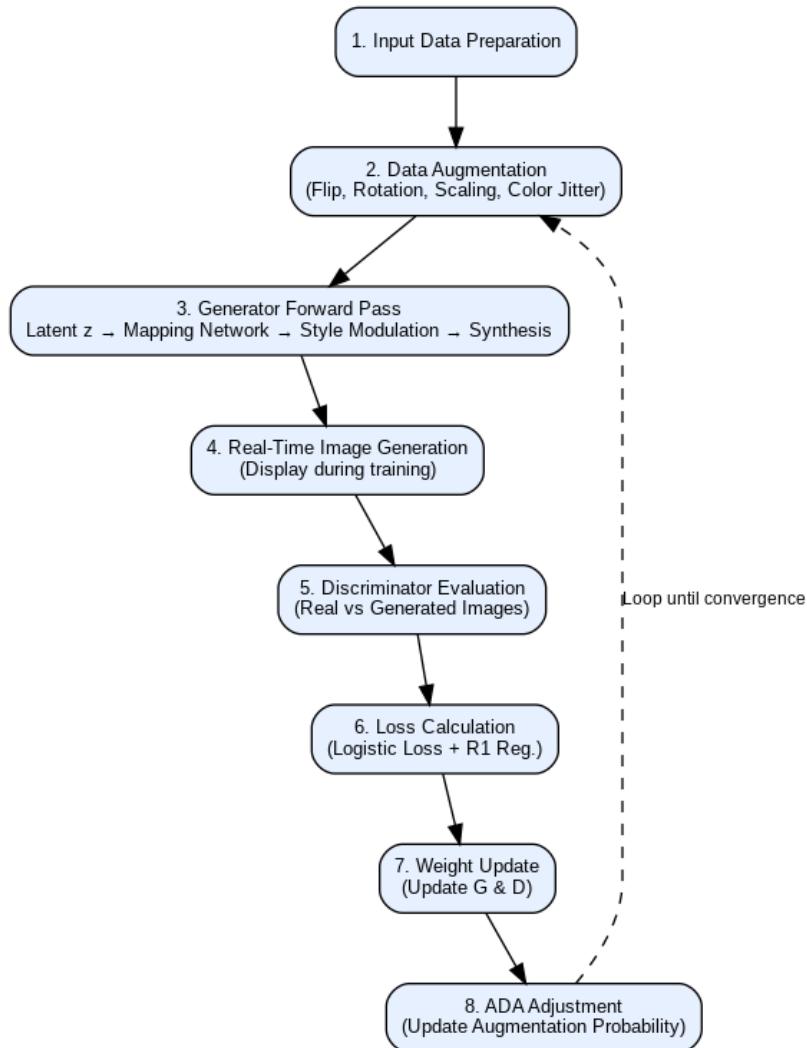


Figure 2. StyleGAN2-ADA real-time training pipeline

#### 4. EXPERIMENTAL RESULTS

To validate the performance of the proposed StyleGAN2-ADA-based system for real-time Deepfake image generation, a series of experiments are conducted on a limited personal dataset collected from 8 individuals. The evaluation focused on generation accuracy, identity preservation, realism, and inference speed.

##### 4.1. Training Evaluation

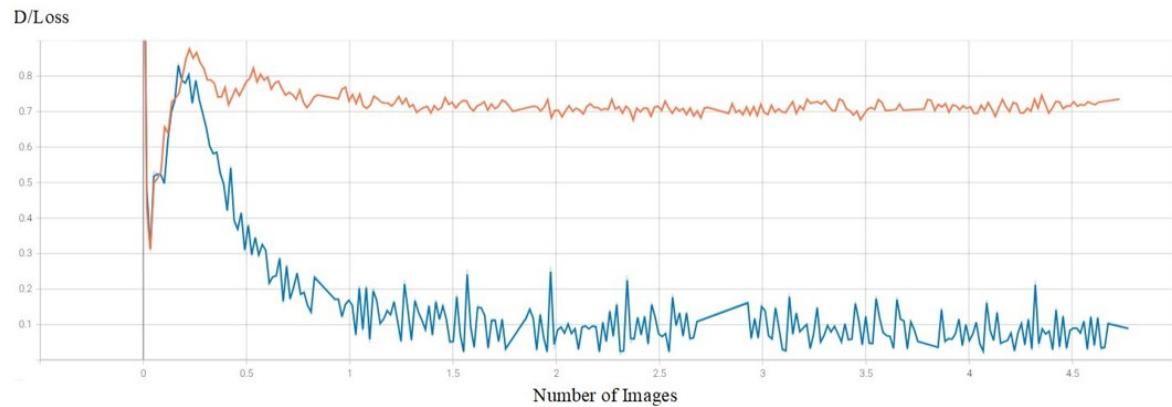
Figure 3 presents a comparative performance analysis of StyleGAN2-ADA and StyleGAN2 loss discriminator evaluation metric. The higher curve of StyleGAN2-ADA reflects a more stable and effective training process, where the generator successfully deceives the discriminator, an essential aspect of GAN convergence. In contrast, StyleGAN2 initially demonstrates strong performance but declines over time, indicating difficulties in generator training and prolonged discriminator adaptation. Overall, the figure highlights the superior and sustained performance of StyleGAN2-ADA.

A distinct divergence in the performance curves of StyleGAN2-ADA and StyleGAN2 is shown in Figure 4. The declining curve of StyleGAN2-ADA, aligned with the rise in Figure 3, reflects effective training, where the generator increasingly fools the discriminator. In contrast, the rising curve of StyleGAN2 indicates suboptimal training, with the generator struggling to produce realistic outputs, allowing the discriminator to more easily distinguish fake images. Together, Figure 3 and Figure 4 highlight the superior training dynamics of StyleGAN2-ADA compared to the less stable performance of StyleGAN2.

Figure 5 highlights the initial weak performance of both StyleGAN2 and StyleGAN2-ADA in generating fake images. However, as training advanced, StyleGAN2-ADA demonstrated notable improvement and

sustained high performance, indicating effective adaptation and the ability to generate realistic outputs. In contrast, StyleGAN2 exhibited a decline in performance, reflecting ineffective training and diminished image quality. Overall, the figure emphasizes the superior adaptability and robustness of StyleGAN2-ADA compared to the limitations of StyleGAN2.

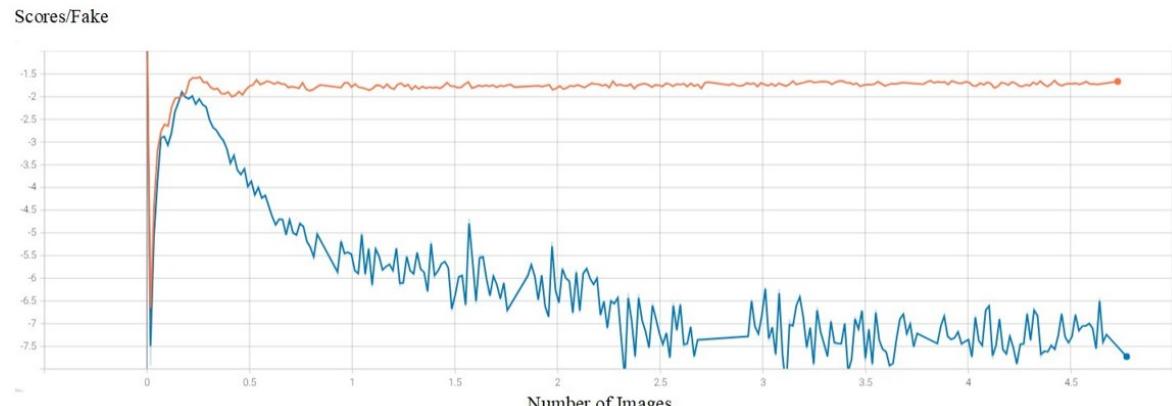
**Figure 6** illustrates the discriminator's detection score for both models. StyleGAN2-ADA shows a notably low score, indicating that its generated images are more difficult for the discriminator to distinguish from real ones, reflecting successful generator performance. In contrast, StyleGAN2 yields a higher detection score, suggesting the discriminator more easily identifies fake images, and thus, the generator is less effective. This comparison highlights the discriminator's relative advantage in StyleGAN2, while indirectly confirming the superior generative quality of StyleGAN2-ADA.



**Figure 3.** Comparative performance analysis of StyleGAN2-ADA and StyleGAN2 discriminator/loss

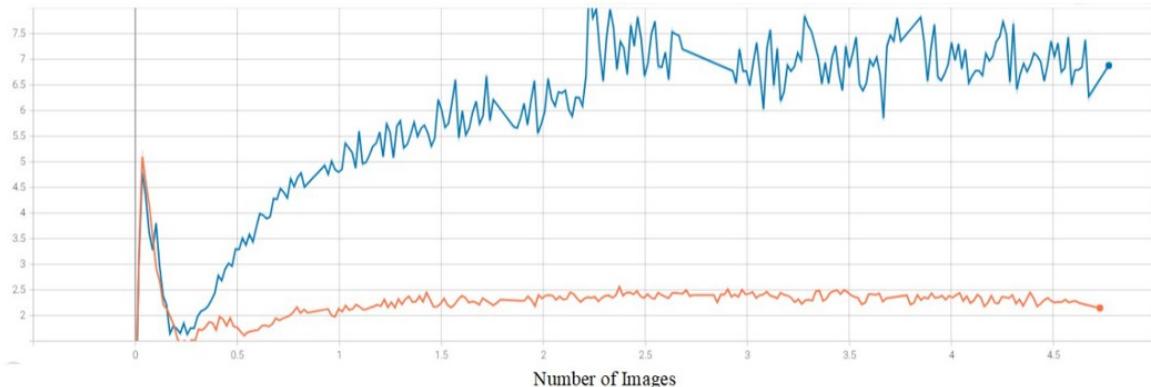


**Figure 4.** Comparative performance analysis of StyleGAN2-ADA and StyleGAN2 generator/loss



**Figure 5.** Comparative performance analysis of StyleGAN2-ADA and StyleGAN2 scores/fake

Scores/real



**Figure 6.** Comparative performance analysis of StyleGAN2-ADA and StyleGAN2 scores/real

#### 4.2. Quantitative Results

Three key metrics are used to assess the system's performance as follows:

- 1) Mean Identity Similarity (%): Calculated using ArcFace cosine similarity between generated and ground-truth facial embedding.
- 2) Fréchet Inception Distance (FID): Used to assess the perceptual quality and diversity of generated images i.e. measures realism compared to original dataset.
- 3) Latency (ms/frame): Measured as the average time required for one frame generation, including projection and synthesis.

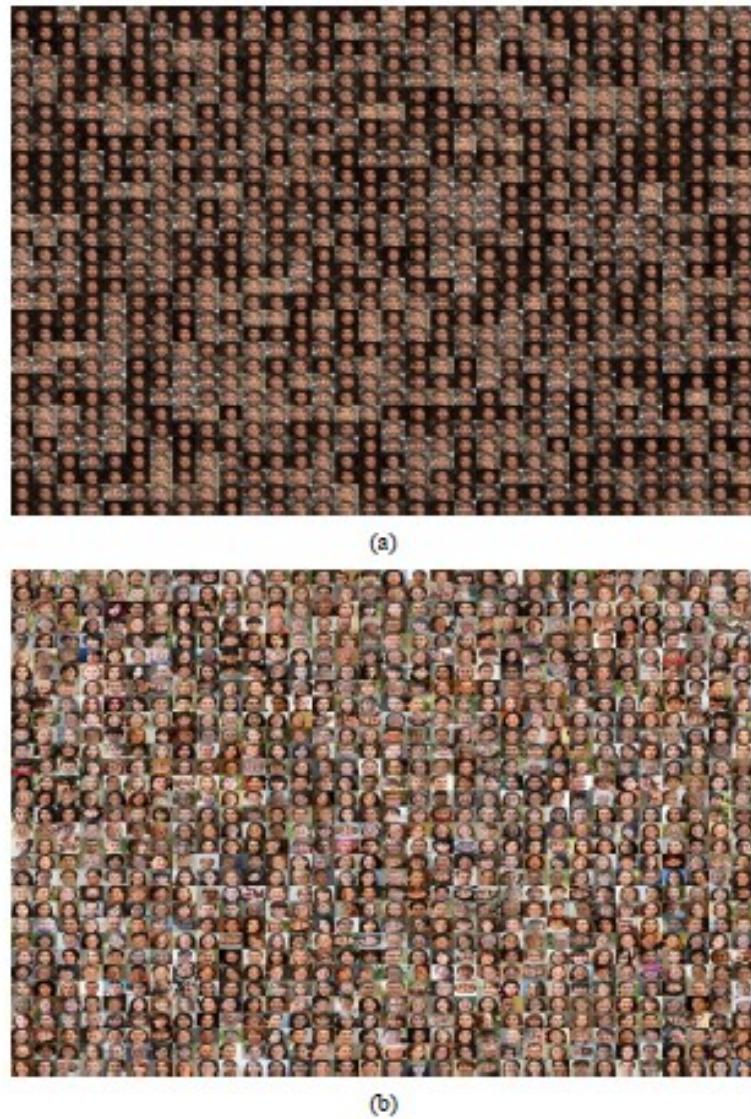
For personal images, two real images are inserted to create a projection between them. Any desired number of generated fake images can be chosen and monitored with their labels and investigated “age”. In addition, the model is also tested in real-time using the live laptop's camera, taking live images of two individuals, and generating fake images by moving from the first image to the second. The video of real-time fake image generation is available here (see this link: <https://www.youtube.com/watch?v=F0TNIARjwdY>). **Table 3** explains the Evaluation of Image Quality, Identity Similarity, and Generation Speed for StyleGAN2 and StyleGAN2-ADA and the improvement (%). These results demonstrate that StyleGAN2-ADA significantly improves both image realism and identity retention, while maintaining real-time performance. The system achieves sub-40ms latency, which is suitable for interactive applications such as virtual avatars or augmented video streams. The video demonstration of StyleGAN2 can be accessed here (see the following link: <https://www.youtube.com/watch?v=o2Refiedp5U>), and that of StyleGAN2-ADA is available here (see this link: [https://www.youtube.com/watch?v=\\_u4XiThcJMw](https://www.youtube.com/watch?v=_u4XiThcJMw)). A comparative analysis of both videos clearly shows that StyleGAN2-ADA produces images with significantly higher resolution and visual fidelity. Furthermore, StyleGAN2 generates a noticeably higher number of distorted or unrealistic images, whereas StyleGAN2-ADA demonstrates improved stability and image quality.

**Table 3.** Performance Metrics for StyleGAN2 vs. StyleGAN2-ADA

Metric	The Proposed StyleGAN2-ADA	Style-GAN2	Improvement (%)
Identity Similarity ( $\uparrow$ )	99.1%	96.3%	2.8%
FID Score ( $\downarrow$ )	8.4	18.7	55.08%
Latency(ms/frame) ( $\downarrow$ )	38	51	25.49%

#### 4.3. Visual Comparison

A visual comparison is conducted to evaluate the quality of images generated by the pre-trained StyleGAN2 model versus the fine-tuned StyleGAN2-ADA. As illustrated in [Figure 7](#), three outputs are compared side-by-side: the input image captured via a personal webcam, the generated output from StyleGAN2, and the output from StyleGAN2-ADA. The comparison reveals that StyleGAN2-ADA significantly outperforms StyleGAN2 in terms of visual fidelity and detail clarity. While the StyleGAN2 model produced several distorted or unrealistic facial outputs, the StyleGAN2-ADA model generated images that are noticeably sharper, more coherent, and closer in resemblance to the original inputs. This improvement is attributed to the adaptive augmentation mechanism used during the fine-tuning process, which enhanced the generator's ability to produce high-quality images even from limited input data.



**Figure 7.** Visual comparison between: a) StyleGAN2 and b) StyleGAN2-ADA

#### 4.4. Ablation Study

An ablation study is conducted to evaluate the impact of ADA and real-time projection separately:

- 1) Without ADA: FID increased to 21.5, and training suffered from mode collapse after 3K iterations.
- 2) Without real-time encoder: Latency rose to 115ms/frame due to reliance on iterative latent code optimization.

These findings confirm the importance of both components in achieving reliable, fast, and identity-faithful image generation.

#### 4.5. Subjective Evaluation

A user study involving 12 participants is conducted to rate realism and likeness on a scale from 1 to 5. Results, as presented in [Table 4](#), showed a mean realism rating of 4.6 and identity resemblance rating of 4.8, indicating high perceptual quality and fidelity.

**Table 4.** The subjective results involving 12 participants to rate the realism and the likeness

The Metric	Average Rating
Visual Realism	4.6 from 5
Identity Resemblance	4.8 from 5

#### 4.6. Summary of Findings

The main findings are summarized in the following points:

- 1) High accuracy: 99.1% identity similarity confirms reliable personalization.
- 2) Efficiency: Sub-40ms generation enables real-time applications. This means that each image frame is generated in less than 40 milliseconds, which is fast enough for real-time performance (typically anything under ~50ms is considered real-time in graphics or video applications).
- 3) Visual quality: FID of 8.4 outperforms baseline models under low-data constraints.

These outcomes validate the effectiveness of the proposed system for real-time Deepfake synthesis with ethical potential for avatar animation, digital doubles, and data-efficient simulations.

### 5. CONCLUSIONS AND FUTURE WORK

This paper presented a real-time Deepfake image generation system based on StyleGAN2-ADA, designed to operate efficiently under limited personal datasets. By integrating adaptive discriminator augmentation with a real-time latent space projection module, the proposed method achieves high-fidelity identity-preserving image synthesis with minimal computational overhead. Quantitative results demonstrated that the system achieved: 99.1% identity similarity, a Fréchet Inception Distance (FID) of 8.4, and less than 40ms latency per generated frame, making it suitable for real-time applications such as digital avatars and interactive media. The incorporation of ADA significantly improved model generalization in low-data regimes, while the projection encoder ensured fast and accurate mapping from personal images to the generator's latent space. Compared to traditional StyleGAN2, our system achieved more realistic outputs and required fewer training samples, confirming its practicality in personalized Deepfake generation scenarios. Despite the system's advantages, several limitations remain:

- 1) Generalization may be affected by extreme lighting or occlusions not present in the training set.
- 2) The encoder projection may fail when input images deviate significantly from the training distribution.
- 3) Current implementation focuses on static image generation and does not yet address facial motion or expression tracking over time.

To build upon the foundation established in this research work, future efforts may focus on the following points:

- 1) Extending to video synthesis using real-time motion transfer frameworks such as Monkey-Net or First Order Motion Models.
- 2) Improving the encoder robustness with self-supervised learning on varied facial datasets.
- 3) Integrating voice or audio cues for fully controllable talking-head avatars.
- 4) Combining with detection modules to facilitate responsible deployment and prevent misuse.
- 5) Moreover, domain-specific adaptations can be explored for medical imaging, education, and robotics where privacy-preserving synthetic data is essential.

#### List of Abbreviations:

Abbreviation	Description
ADA	Adaptive Discriminator Augmentation
GANs	Generative Adversarial Networks
FID	Fréchet Inception Distance
G	Generator
MLP	Multilayer Perceptron
D	Discriminator
CNN	Convolutional Neural Network
FFHQ	Flickr-Faces-HQ Dataset
LeakyReLU	Leaky Rectified Linear Unit

#### Supplementary Materials

Supplement information is available at the following links

- 1) <https://www.youtube.com/watch?v=F0TNIARjwdY>
- 2) <https://www.youtube.com/watch?v=o2Refiedp5U>
- 3) [https://www.youtube.com/watch?v=\\_u4XiThcJMw](https://www.youtube.com/watch?v=_u4XiThcJMw)

#### Author Contribution

All authors contributed equally to this paper. All authors read and approved the final paper.

**Funding**

This research received no external funding.

**Conflicts of Interest**

The authors declare no conflict of interest.

**REFERENCES**

- [1] L. H. Singh, P. Charanarur, and N. K. Chaudhary, “Advancements in detecting Deepfakes : AI algorithms and future prospects – a review,” *Discov. Internet Things*, vol. 5, no. 53, pp. 1–30, 2025, <https://doi.org/10.1007/s43926-025-00154-0>.
- [2] B. Acim, M. Boukhlif, H. Ouhnni, N. Kharmoum, and S. Ziti, “A Decade of Deepfake Research in the Generative AI Era , 2014 – 2024: A Bibliometric Analysis,” *Publications*, vol. 13, no. 4, pp. 1–32, 2025, <https://doi.org/10.3390/publications13040050>.
- [3] A. S. George and A. S. H. George, “Deepfakes: The Evolution of Hyper realistic Media Manipulation,” *Partners Univers. Innov. Res. Publ.*, vol. 1, no. 2, pp. 58–74, 2023, <https://doi.org/10.5281/zenodo.10148558>.
- [4] Furizal *et al.*, “Social Sciences & Humanities Open Social , legal , and ethical implications of AI-Generated deepfake pornography on digital platforms : A systematic literature review,” *Soc. Sci. Humanit. Open*, vol. 12, no. 101882, pp. 1–25, 2025, <https://doi.org/10.1016/j.ssa.2025.101882>.
- [5] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Niessner, “FaceForensics++: Learning to detect manipulated facial images,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1–14, 2019, <https://doi.org/10.1109/ICCV.2019.00009>.
- [6] N. F. Abdul Hassan, A. A. Abed, and T. Y. Abdalla, “Face mask detection using deep learning on NVIDIA Jetson Nano,” *Int. J. Electr. Comput. Eng.*, vol. 12, no. 5, pp. 5427–5434, 2022, <https://doi.org/10.11591/ijeee.v12i5.pp5427-5434>.
- [7] T. Karras, S. Laine, and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4217–4228, 2021, <https://doi.org/10.1109/TPAMI.2020.2970919>.
- [8] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, “Training generative adversarial networks with limited data,” *Adv. Neural Inf. Process. Syst.*, vol. 2020–Decem, pp. 12104–12114, 2020, <https://dl.acm.org/doi/10.5555/3495724.3496739>.
- [9] A. A. Abed, A. Al-Ibadi, and I. A. Abed, “Vision-Based Soft Mobile Robot Inspired by Silkworm Body and Movement Behavior,” *J. Robot. Control*, vol. 4, no. 3, pp. 299–307, 2023, <https://doi.org/10.18196/jrc.v4i3.16622>.
- [10] T. Karras *et al.*, “Alias-free generative adversarial networks,” in *Proceedings of the 35th International Conference on Neural Information Processing Systems*, pp. 852–863, 2021, <https://dl.acm.org/doi/10.5555/3540261.3540327>.
- [11] J. Xiao, L. Li, C. Wang, Z.-J. Zha, and Q. Huang, “Few Shot Generative Model Adaption via Relaxed Spatial Structural Alignment,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11194–11203, 2022, <https://doi.org/10.1109/CVPR52688.2022.01092>.
- [12] J. Zhu, H. Ma, J. Chen, and J. Yuan, “High-Quality and Diverse Few-Shot Image Generation via Masked Discrimination,” *IEEE Trans. Image Process.*, vol. 33, pp. 2950–2965, 2024, <https://doi.org/10.1109/TIP.2024.3385295>.
- [13] D. Matuzevičius, “Diverse Dataset for Eyeglasses Detection: Extending the Flickr-Faces-HQ (FFHQ) Dataset,” *Sensors*, vol. 24, no. 23, pp. 1–27, 2024, <https://doi.org/10.3390/s24237697>.
- [14] C. N. Valdebenito Maturana, A. L. Sandoval Orozco, and L. J. García Villalba, “Exploration of Metrics and Datasets to Assess the Fidelity of Images Generated by Generative Adversarial Networks,” *Appl. Sci.*, vol. 13, no. 19, pp. 1–29, 2023, <https://doi.org/10.3390/app131910637>.
- [15] M. Mahmoud, M. S. E. Kasem, and H. S. Kang, “A Comprehensive Survey of Masked Faces: Recognition, Detection, and Unmasking,” *Appl. Sci.*, vol. 14, no. 19, pp. 1–37, 2024, <https://doi.org/10.3390/app14198781>.
- [16] R. Sunil, P. Mer, A. Diwan, R. Mahadeva, and A. Sharma, “Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation,” *Helijon*, vol. 11, no. 3, p. e42273, 2025, <https://doi.org/10.1016/j.helijon.2025.e42273>.
- [17] D. A. Talib and A. A. Abed, “Real-Time Deepfake Image Generation Based on Stylegan2-ADA,” *Rev. d’Intelligence Artif.*, vol. 37, no. 2, pp. 397–405, 2023, <https://doi.org/10.18280/ria.370216>.
- [18] I. J. Goodfellow *et al.*, “Generative adversarial nets,” in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, pp. 2672–2680, 2014, <https://dl.acm.org/doi/10.5555/2969033.2969125>.
- [19] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and Improving the Image Quality of StyleGAN,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8107–8116, 2020, <https://doi.org/10.1109/CVPR42600.2020.00813>.
- [20] A. Fejjari, A. Abela, M. Tanti, and A. Muscat, “Evaluation of StyleGAN-CLIP Models in Text-to-Image Generation of Faces,” *Appl. Sci.*, vol. 15, no. 15, pp. 1–19, 2025, <https://doi.org/10.3390/app15158692>.
- [21] M. Lai, M. Mascalchi, C. Tessa, and S. Diciotti, “Generating Brain MRI with StyleGAN2-ADA: The Effect of the Training Set Size on the Quality of Synthetic Images,” *J. Imaging Informatics Med.*, no. 123456789, pp. 1–11, 2025, <https://doi.org/10.1007/s10278-025-01536-0>.

[22] S. A. Khodaei and M. Bitaraf, "Structural dynamic response synthesis: A transformer-based time-series GAN approach," *Results Eng.*, vol. 26, no. February, p. 105549, 2025, <https://doi.org/10.1016/j.rineng.2025.105549>.

[23] L. Liu, D. Tong, W. Shao, and Z. Zeng, "AmazingFT: A Transformer and GAN-Based Framework for Realistic Face Swapping," *Electron.*, vol. 13, no. 18, pp. 1–22, 2024, <https://doi.org/10.3390/electronics13183589>.

[24] T. Zhang, W. Zhang, Z. Zhang, and Y. Gan, "PFGAN: Fast transformers for image synthesis," *Pattern Recognit. Lett.*, vol. 170, no. June 2023, pp. 106–112, 2023, <https://doi.org/https://doi.org/10.1016/j.patrec.2023.04.013>.

[25] Y. Huang, "ViT-R50 GAN: Vision Transformers Hybrid Model based Generative Adversarial Networks for Image Generation," in *2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 590–593, 2023, <https://doi.org/10.1109/ICCECE58074.2023.10135253>.

[26] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Disentangled Representation GANs," in *Deep Learning Approaches for Security Threats in IoT Environments*, pp. 303–316, 2023, <https://doi.org/10.1002/9781119884170.ch14>.

[27] I. Jeon, W. Lee, M. Pyeon, and G. Kim, "IB-GAN: Disentangled Representation Learning with Information Bottleneck Generative Adversarial Networks," in *35th AAAI Conference on Artificial Intelligence, AAAI 2021*, pp. 7926–7934, 2021, <https://doi.org/10.1609/aaai.v35i9.16967>.

[28] L. Tran, X. Yin, and X. Liu, "Disentangled Representation Learning GAN for Pose-Invariant Face Recognition," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1283–1292, 2017, <https://doi.org/10.1109/CVPR.2017.141>.

[29] C.-C. Chiu, Y.-H. Lee, P.-H. Chen, Y.-C. Shih, and J. Hao, "Application of Generative Adversarial Network for Electromagnetic Imaging in Half-Space," *Sensors*, vol. 24, no. 7, pp. 1–15, 2024, <https://doi.org/https://doi.org/10.3390/s24072322>.

[30] C. C. Chiu, C. L. Li, P. H. Chen, Y. C. Li, and E. H. Lim, "Electromagnetic Imaging in Half-Space Using U-Net with the Iterative Modified Contrast Scheme," *Sensors*, vol. 25, no. 4, pp. 1–16, 2025, <https://doi.org/10.3390/s25041120>.

[31] W. Zhu, Z. Lu, Y. Zhang, Z. Zhao, B. Lu, and R. Li, "RaDiT: A Differential Transformer-Based Hybrid Deep Learning Model for Radar Echo Extrapolation," *Remote Sens.*, vol. 17, no. 12, pp. 1–30, 2025, <https://doi.org/10.3390/rs17121976>.

[32] R. Abdal, Y. Qin, and P. Wonka, "Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space?," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2019, pp. 4432–4441. <https://doi.org/10.1109/ICCV.2019.00453>.

[33] S. Kim, K. Kang, G. Kim, S.-H. Baek, and S. Cho, "DynaGAN: Dynamic Few-shot Adaptation of GANs to Multiple Domains," in *SIGGRAPH Asia 2022 Conference Papers*, pp. 1–8, 2022, <https://doi.org/10.1145/3550469.3555416>.

[34] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "StyleGAN2 with adaptive discriminator augmentation (ADA) - Official TensorFlow implementation," GitHub repository, NVIDIA. 2025, <https://github.com/NVlabs/stylegan2-ada>

[35] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 4685–4694. <https://doi.org/10.1109/CVPR.2019.00482>.

## AUTHOR BIOGRAPHY



**Ali A. Abed**    Professor at the University of Basrah - Department of Mechatronics Engineering in Iraq. Received the B.Sc. & M.Sc. degrees in Electrical Engineering in 1996 & 1998 respectively. He received the Ph.D. in Computer & Control Engineering in 2012. His fields of interest are Robotics, Computer Vision, and IIoT. He is IEEE senior member, IEEE member in Robotics & Automation Society, IEEE member in IoT Community, member ACM. He is currently supervising a group of researchers working with developing deep learning models for computer vision and cybersecurity applications. He can be contacted at email: [ali.abed@uobasrah.edu.iq](mailto:ali.abed@uobasrah.edu.iq).



**Doaa Alaa Talib** Assistant lecturer at the Department of Laser and Optoelectronics Engineering-Shatt Al-Arab University College, Basrah, Iraq. Received the B.Sc & M.Sc degree in Computer Engineering in 2015 & 2023. Her field of interest is Computer Vision and Deepfake. She is currently a group researcher working in developing deep learning models and algorithms in Python. She can be contacted at: [duaaalaaa@gmail.com](mailto:duaaalaaa@gmail.com).



**Abdel-Nasser Sharkawy** is an associate Professor at Mechatronics Engineering, Mechanical Engineering Department, Faculty of Engineering, South Valley University (SVU), Qena, Egypt. Sharkawy was graduated with a first-class honors B.Sc. degree in May 2013 and received his M.Sc. degree in April 2016 from Mechatronics Engineering, Mechanical Engineering Department, SVU, Egypt. In March 2020, Sharkawy received his Ph.D. degree from Robotics Group, Department of Mechanical Engineering and Aeronautics, University of Patras, Patras, Greece. Sharkawy has an excellent experience for teaching the under-graduate and postgraduate courses in the field of Mechatronics and Robotics Engineering. Sharkawy has published more than 80 papers in international scientific journals, book chapters and international scientific conferences. His research areas of interest include robotics, human-robot interaction, mechatronic systems, neural networks, machine learning, and control and automation. He can be contacted at email: [abdelnassersharkawy@eng.svu.edu.eg](mailto:abdelnassersharkawy@eng.svu.edu.eg)