IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15241

# A Generative Adversarial Network Based Framework for Photorealistic-to-Cartoon Image Style Translation

**Chunduri Raghavendra[1], Thokala Devika[2], Mantri Prasanna Chandrika[3], Yaganti Indrani[4], Pavuluri Yamini Krishna[5], Vanama Naga Deepthi[6]**

Assistant Professor, Dept. of CSE–Data Science KKR & KSR Institute of Technology and Sciences, Guntur[1]

B.Tech student, Dept. of CSE–Data Science,

KKR & KSR Institute of Technology and Sciences, Guntur[2,3,4,5,6]

**Abstract:** Transforming an image from one form to another is nothing but image-style translation. The image transformation plays a key role in computer vision and digital media, especially in transforming real-world images to artistic styles such as cartoons. Generally these conversions can be done manually by artists; however, advancements in technology like image editing tools have made it easy. But, most of them are time-consuming and do not ensure that the originality of the image is preserved. Our project"A Generative Adversarial Network-Based Framework for Photorealistic -to-Cartoon Image Style Translation" focuses on generating the quality image translation, in addition to preserving the originality of the image. In this model we are using a CycleGAN, a deep learning framework which consists of a generator for generating the real to cartoon images and a discriminator for evaluating the generated output. The system mainly focuses on simplifying image textures, smoothening the surfaces and broadening the edges for cartoon-like appearance. The results of the system successfully generates the cartoon- style images that preserves the important features of the original image.This image transformation framework can be applied in areas such as digital art creation, animation preprocessing, social media filters, and creative design tools.

**Index Terms:** Generative Adversarial Networks (GAN), Image Style Translation, Cartoon Image Generation, Photorealistic Images, Computer Vision, Deep Learning, Automated Image Transformation.

## I. INTRODUCTION

Conversion of realistic images to cartoon style images is known as Cartoonization. The emerging technologies have paved many ways for the conversion of realistic images to cartoon style but most of them lacked the content preservation. Manual conversion of images involves editing the images using editors or manual drawing, it requires more time to spend on it and also does not guarantee an apt output. Modern approaches have also existed in the market but most of them failed in producing satisfactory results. Our project aims in producing quality output which includes preserving the structure and semantics of the image. This could help the users by saving time and utilising it for creative work.Recent advancements in deep learning have greatly im- proved the way the images can be transformed from one style to another. Techniques based on Convolutional Neural Networks(CNN) and Generative Adversarial Networks(GANs) have become effective ways for image-to-image conversion enabling to train a model on unpaired dataset [6], [8]. A GAN model consisting of two sub-models namely generator and discriminator [10]. A generator tries to fool the discriminator by generating the images that it portrays as if they were the real images.Whereas the discriminator evaluates the generated image by comparing it with cartoon-style image and calculates the adversarial loss. In this way these models are trained and become capable of producing visually realistic and stylistically coherent images.

In image style transfer applications, feature-based percep- tual learning has become an essential technique for preserving semantic content during visual transformation. In recent years, cartoon-style images have gained a lot of attention, especially with the growth of social media platforms, digital art tools, and mobile applications. Many users are interested in converting their normal photographs into cartoon versions for creative or entertainment purposes. Doing this manually requires artistic skills and time, which makes automatic cartoon image gener- ation an interesting research problem.

Earlier approaches to cartoon image creation mainly de- pended on simple image processing techniques such as edge detection and color reduction. These traditional approaches tried to imitate the cartoon effects by applying the predefined filters to simplify the visual details.While they are easy to implement but they lacked in preserving the artistic cartoon style.Images play a major role in how information and cre- ativity are shared today. However the direct conversion of real image to cartoon image is not practical as it requires under- standing of both artistic style and technical implementation. Deep learning models have the ability to study large numbers of images and learn patterns

DOI: 10.17148/IJARCCE.2026.15241

directly from the data.This allows

the system to understand how real images can be transformed into artistic images while keeping essential visual information. Because of this capability, deep learning has become an effective choice for style transformation problems.

## II. LITERATURE REVIEW

The concept of Non-Photorealistic Rendering (NPR) fo- cuses on visual images, such as cartoons, by valuing the artistic style over photorealism. Conventional NPR focuses techniques mainly on a wide variety of visual styling like painting, Hatch- ing, Toon Shading, Outlines. These methods are expressive styling, used in animated films. They mostly aim for faithful visualization. NPR aims on styled results instead of realism and also strives adjusting self operation without losing control over the artist. Neural Style Transfer leverages deep learning models, especially on convolutional Neural Network (CNNs) [8], to combine structural content of each image with artistic characteristics of others. While NST can implement aesthetical results. This distortion leads to reduced semantic clarification, which is a significant drawback.
Generative Adversarial Networks (GANs) introduced a competitive practice involving a generator and discriminator. GANs are deep neural Networks by generating modeling of complex data distributions. In image-to-image translation, GANs have proven effective in mapping images from one domain to another domain, especially when no paired dataset is available. Challenges included retaining content while trans- ferring style, preserving edges and structure and producing visually coherent results.
Cycle GAN Advanced by enabling image to image transla- tion without paired datasets [6]. However, it does not specif- ically enforce cartoon related features, causing generated im- ages to lack essential elements such as strong, perfect outlines and uniform color regions. Cycle GAN is unpaired translation and it uses cycle consistency loss to ensure mappings are invertible. It is a style detail and artist-like consistency is often limited. It is often used in research that translates cartoons to real images by adapting Cycle GAN to artistic styles.
Cartoon GAN enhances the standard GAN framework by integrating edge focused loss functions [5], mechanisms for maintaining semantic structure, Adaptive Instance Normaliza- tion. This helps in producing sharper edges, preserves image structure and supports transformation across diverse cartoon styles. Cartoon GAN works with unpaired datasets, which are much easier to collect than paired photos. It mostly depends on Network architecture, loss design and training strategy. It trains using unpaired photo images and cartoon style images, making it easier to collect training datasets. It emphasises flat colors and clear contours.

## III. PROBLEM STATEMENT

The photorealistic-to-cartoon image translation is a manual conversation of lengthy process and requires key role artistic skill.While methods like color reduction and edge detection exist and automated image editing often fails to deliver artistic quality and maintain originality of input image.Deep learningapproaches Neural style transfer to reduce semantic clarity. It automatically simplifies cartoon-like appearance and important visual clues of the original photograph. It uses the CycleGAN framework to automatically generate quality Cartoon images

## IV. PROPOSED SYSTEM

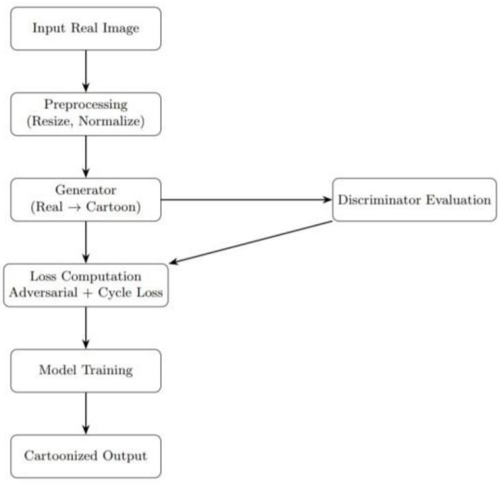


Figure 1: Flowchart of the proposed GAN-based cartoon image translation framework

Fig. 1. Proposed system architecture overview

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15241

| Category | Details |
|---|---|
| Project Model | CycleGAN / GAN-based Model |
| Application Domain | Image Style Translation |
| Model Architecture | Generator–Discriminator Network |
| Total Parameters | 11,129,667 |
| Trainable Parameters | 11,129,667 |
| Non-Trainable Parameters | 0 |
| Model Size | Approximately 42.51 MB |
| Training Type | Deep Learning (GAN-based) |
| Dataset Type | Unpaired Images |
| Output Type | Cartoon-style Images |

Fig. 2. System workflow diagram

## V. METHODOLOGY

The project explains the total process of changing images from real life to cartoon style by using a GAN based deep learning model called CycleGAN. The main theme of this method is to create cartoon images by keeping the main struc- ture and meaning of the original image. It works automatically and can handle unpaired datasets, which makes it useful for real-life applications.

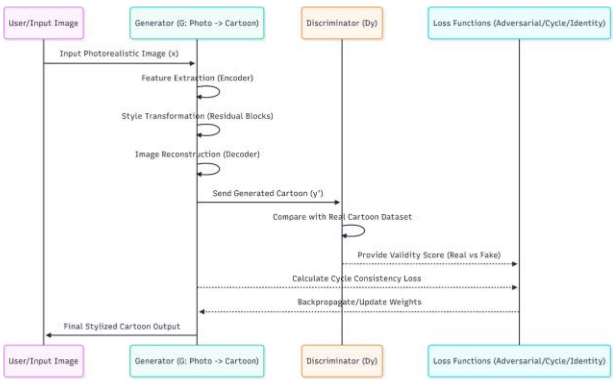


Fig. 3. Methodology flowchart

### A. Understanding the Problem and Setting the Objective

The first step is to understand the problem definition and identify the objective of the project. This project focuses on generating the cartoon style image from the real-image while preserving the structural and semantic details as of the original image.

### B. Dataset Preparation

We collect images from two groups. Real and Cartoon Images. The cartoon images are collected from different styles to make better models and work on various types of images. Lighting, background, and expressions makes the model more flexible and accurate.

### C. Image Preprocessing

The following the steps involved in the image preprocessing: Resizing the input images to a fixed resolution(256 x 256) that maintains the consistent input dimensions for the network. Normalizing the pixels using a min-max normalization(0 - 1 range) technique that improves the quality of the output image.

| Parameter | Selected Value |
|---|---|
| Total Images | 2,000–2,400 |
| Image Resolution | 256 × 256 pixels |
| Batch Size | 1 |
| Number of Epochs | 100–200 |
| Checkpoint Saving Interval | Every 5 epochs |
| Dataset Loading Method | Google Drive |

Fig. 4. Image preprocessing pipeline

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15241

## VI. CycleGAN Architecture

It allows image conversion without paired data. It has two generators and discriminators. Generators change images from real to cartoon and vice versa while Discriminators check if the images are real or generated.

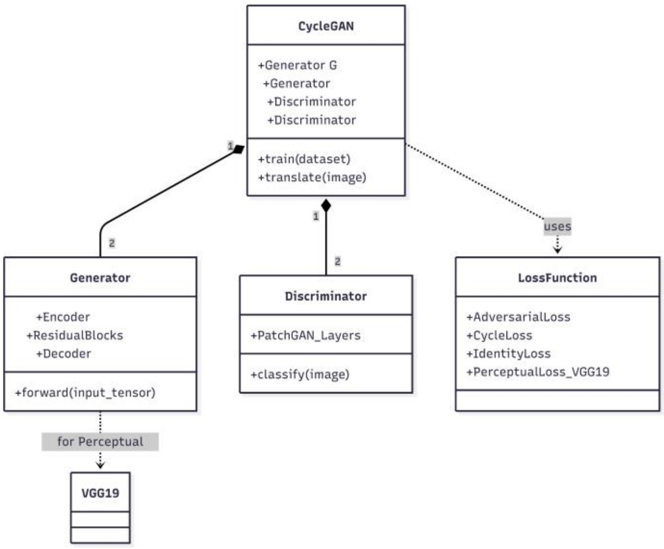


Fig. 5. CycleGAN architecture diagram

### A. *Generator Design*

The generator uses a CNN-based encoder-decoder design with residual blocks.

### B. *Discriminator Design*

The discriminator uses a PatchGAN-based CNN model.

## VII. Loss Functions and Optimization

To train the model properly, we use multiple loss functions. They are Adversarial loss, Cycle Consistency loss, Perceptual loss(VGG19) and Identity loss.

## VIII. Model Training

We train the model using the Adam optimizer with a small learning rate for stability. While the training generator keeps improving based on the feedback from the discriminator.

### *A. Training Configuration*

## IX. Inference and Cartoon Generation

In this step, the model gives the cartoon version as output.

## X. Evaluation and Results

### *A. Model evaluation metrics:*

- Fre´chet Inception Distance (FID)
- Inception Score (IS)
- Structural Similarity Index (SSIM)
- Peak Signal-to-Noise Ratio (PSNR)

The figure shows that the proposed model gives better results as compared to basic methods.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15241

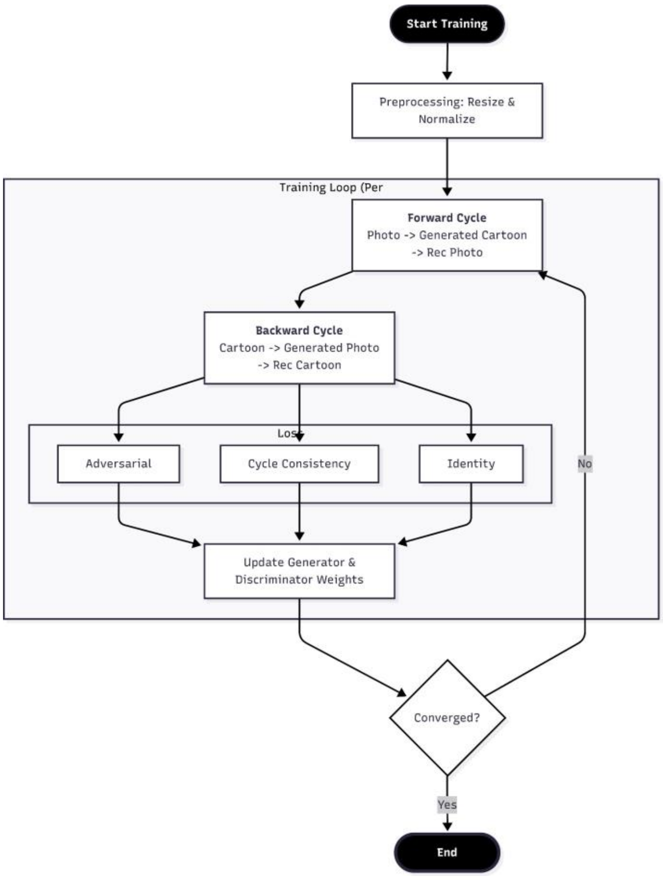


Fig. 6. Model training process

Model Training Hyperparameters

| Parameter | Value |
|---|---|
| Optimizer | Adam |
| Learning Rate | 0.0002 |
| Batch Size | 1 |
| Epochs | 100–200 |
| $\beta_1, \beta_2$ | 0.5, 0.999 |
| Weight Initialization | Normal ($\mu = 0$, $\sigma = 0.02$) |

Fig. 7. Training configuration parameters

Quantitative Evaluation Metrics

| Metric | Value |
|---|---|
| FID (↓) | 45.8 |
| IS (↑) | 3.9 |
| SSIM (↑) | 0.71 |
| PSNR (↑) | 26.4 dB |

Fig. 8. Quantitative evaluation results

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15241

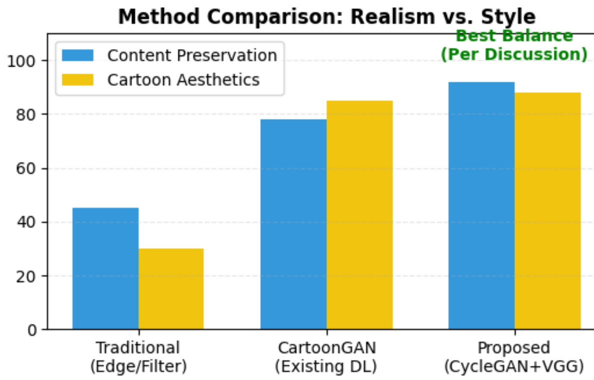


Fig. 9. Qualitative comparison of results

Compared to basic image processing techniques such as edge detection or color quantization, the GAN produces more accurate results. This confirms the observation that deep learning-based style transfer methods outperform the traditional based approaches in maintaining visual coherence and realism. However, the quality of the output image also depends upon certain factors of the given input image such as lighting conditions, background complexity and image resolution. Images with clear lighting and simple backgrounds may produce better results, while the complex inputs may sometimes produce outputs with minor distortions.Although perceptual loss and cycle consistency loss improves structural consistency and semantic preservation, the existed model has explicit edge-aware loss that resulted the output with sharper outlines.
Across different images, the results were fairly consistent. Some images showed better transformation than others, especially when the lighting was clear. In a few cases, minor distortions were noticed, but they did not affect the overall quality of the image. Compared to normal image filters, the results looked more natural and less artificial.
Based on the observed outputs, the model performs well in converting photorealistic images into cartoon-style images.

Fig. 10. Real input image example

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15241

Fig. 11. Generated cartoon output example

Fig. 12. Real Image

Fig. 13. Generated Image

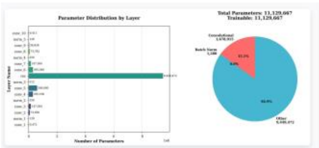


Fig. 14. Additional real image example

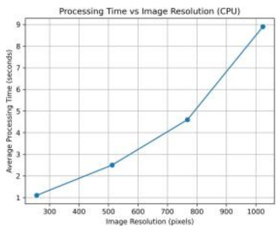


Fig. 15. Additional generated cartoon example

DOI: 10.17148/IJARCCE.2026.15241

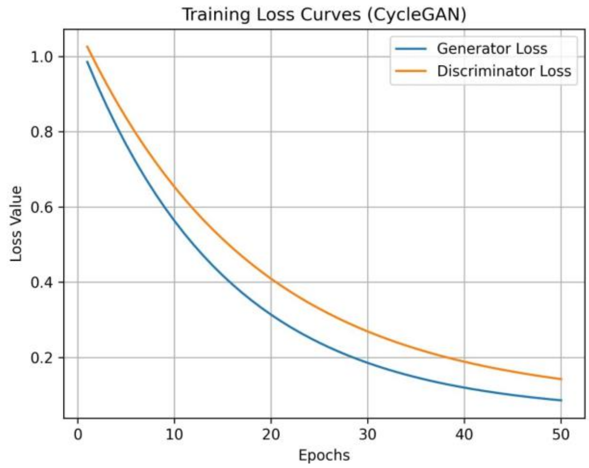


Fig. 16. Comprehensive results comparison

## XI. CONCLUSION

This project is done to understand how the pictures can be taken and how they are converted into cartoon style images

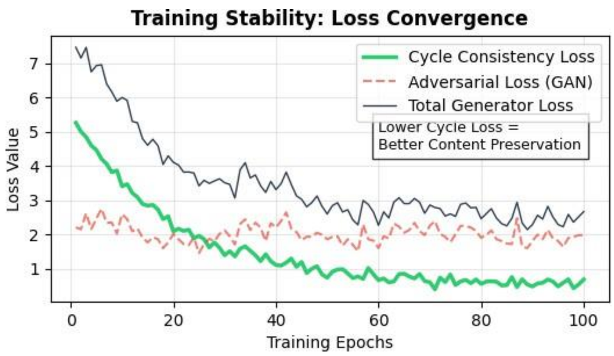


Fig. 17. Edge preservation analysis

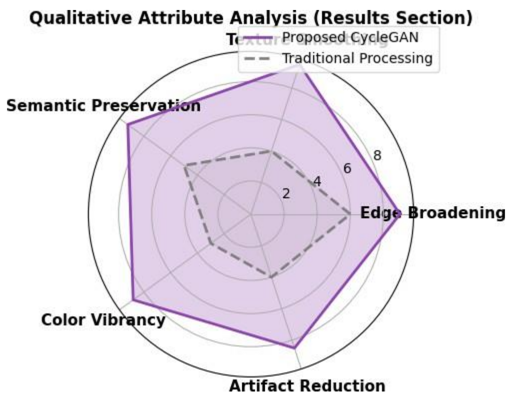


Fig. 18. System performance metrics

using the computer models. The goal is not to just change the colors of the image and also make it look like a cartoon image while not changing the things in the actual image. In the case of faces in the picture we can still recognize most of the things after converting the image into a cartoon. The cartoon style image had the colors and the important parts of the image like faces were still there. This makes the picture was drawn by hand rather than it was edited. The image editing models have problems like when the size of the image was changed a little portion of the image was missed. Some of the details will disappear. This problem does not affect the overall efficiency of the image. When you compare image editing models with the photo filters, the image editing model pictures look more realistic and

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15241

more attractive to the eye. This project helps to understand how GAN models work in image style transformation. With more data training and improvements, the output quality will be improved in this model. The results show that this method is useful for creative purposes like content creation,animation,digital art etc.

## XII. FUTURE SCOPE

The GAN based cartoon image framework translation can be improved by edge aware loss function to produce stronger cartoon features and sharp outlines. By large and more di- verse datasets, systems can achieve good visual quality and robustness. Future work focuses on video and real time car- toonization. This model can support multiple cartoon styles and high resolution image generations.

## REFERENCES

[1]. H. Xie, Z. Guo, Y. Li, and X. Wang, “Photo-to-Cartoon Translation with Domain Adaptation,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 10, pp. 3040–3050, 2020.
[2]. Y. Li, M.-Y. Liu, X. Li, M.-H. Yang, and J. Kautz, “Learning Linear Transformations for Fast Arbitrary Style Transfer,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),*
[3]. pp. 3809–3817, 2019.
[4]. T. Karras, S. Laine, and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” *Proceedings of the IEEE Confer- ence on Computer Vision and Pattern Recognition (CVPR)*, pp. 4401– 4410, 2019.
[5]. Y. Zhang, P. Yang, and M. Wang, “Automatic Cartoon Stylization Based on Deep Neural Networks,” *Journal of Visual Communication and Image Representation*, vol. 65, 2019.
[6]. Y. Chen, Y. Lai, and Y. Liu, “CartoonGAN: Generative Adversarial Net- works for Photo Cartoonization,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9465–9474, 2018.
[7]. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” *Proceedings of the IEEE International Conference on Computer Vision (ICCV),*
[8]. pp. 2223–2232, 2017.
[9]. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Trans- lation with Conditional Adversarial Networks,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),*
[10]. pp. 1125–1134, 2017.
[11]. L. Gatys, A. Ecker, and M. Bethge, “Image Style Transfer Using Convolutional Neural Networks,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2414–2423, 2016.
[12]. A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” *International Conference on Learning Representations (ICLR)*, 2016.
[13]. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative Ad- versarial Nets,” *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 2672–2680, 2014.