iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59168

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

Cartoonization of Images Using Generative Adversarial Network

Haseena Syed[1], Ratnamala Rayi[2], Prabhudev Tamanam[3], Sai Tharun Tata[4], Vijayananda Ratnam. Ch[5]
[1, 2, 3, 4]*UG Students,* [5]*Assistant Professor Department of CSE, Vasireddy Venkatadri Institute of Technology (Autonomous), Guntur, AP*

Abstract: ***This research paper introduces a web-based image style conversion system powered by a trained AnimeGANv2 model. The system enables real-time conversion of images into anime-style artwork within a web application. Through the integration of TensorFlow, Protobuf, and ONNX formats, the AnimeGANv2 model is seamlessly deployed for efficient style conversion. The web application incorporates user authentication, secure image uploading, and instant style conversion features. Evaluation through user testing assesses the system's usability, performance, and user satisfaction. Results demonstrate the successful implementation of real- time image style conversion, offering a practical solution for users interested in artistic rendering. This research contributes to understanding GAN-based image style conversion applications and their integration into web platforms for creative expression and visual content generation.***

Keywords: ***AnimeGANv2, image style conversion, web application, TensorFlow, Protobuf, ONNX, real-time, user authentication, secure image uploading, user satisfaction, GAN-based creative expression, visual content generation***

I. INTRODUCTION

Real-time image-to-anime conversion stands as a significant innovation in digital artistry and creative expression. The emergence of deep learning modes like AnimeGANv2 has enabled users to swiftly transform real-world photographs into compelling anime-style artwork. This transformation not only echoes traditional anime aesthetics but also broadens the horizons for artistic exploration and visual storytelling. Our project centers on harnessing AnimeGANv2 within a web-based platform, providing users with an intuitive interface to convert images to anime versions in real time. Through the integration of TensorFlow, Protobuf, and ONNX formats, we've developed a system that facilitates secure image uploading and swift style conversion. Users can directly upload their photographs to the web application and witness the transformation unfold, experiencing the enchantment of anime- style rendering within moments. The significance of real-time image-to-anime conversion extends beyond technological advancement; it serves to inspire creativity and kindle imagination across various domains. By democratizing access to sophisticated image processing techniques, our project empowers users to delve into novel realms of visual expression and storytelling. Whether for personal enjoyment, artistic exploration, or professional endeavors, the capability to generate anime- style artwork in real time proves invaluable to creators across diverse sectors and industries.

II. LITERATURE SURVEY

1) Chao Dong's [1] study aims to enhance the Super-Resolution Convolutional Neural Network (SRCNN) by proposing a compact hourglass-shaped CNN structure, achieving over 40 times faster processing while maintaining superior restoration quality. Key modifications include the addition of a deconvolution layer, feature dimension reduction, and the adoption of smaller filter sizes with more mapping layers, along with parameter adjustments for real-time performance on standard CPUs.

2) This author, B. Lim [2], introduces the Enhanced Deep Super-Resolution Network (EDSR), leveraging residual learning techniques to achieve superior performance by optimizing model size, stabilizing the training process, and introducing a Multi- Scale Deep Super Resolution system (MDSR), showcasing its excellence through the NTIRE2017 Super Resolution challenges.

3) Yang Chen [3] proposed the CartoonGAN Model, which presents a novel generative adversarial network framework tailored for converting real-world scene photos into high-quality cartoon-style images, outperforming existing methods through its unique combination of semantic content loss, edge-preserving adversarial loss, and efficient training strategies.

4) Author Y. Chen [4] proposes a novel approach inspired by Gatys's work to transform photos into comics using deep convolutional neural networks (CNNs). Addressing the limitations of Gatys's method in minimalist styles like comics, the paper introduces a dedicated comic style CNN trained for classifying comic images and photos, ensuring better comic stylization results. Additionally, the paper modifies the optimization framework to guarantee grayscale output, improving upon the state-of-the-art.

©IJRASET: All Rights are Reserved

III. RELATED WORK

A. Non-Photorealistic Rendering (NPR) in Anime Style Conversion

Non-photorealistic Rendering (NPR) in anime-style conversion involves the transformation of real-world images into anime-style art works using sophisticated deep learning techniques and stylization algorithms. In our project, AnimeGANv2 stands as the cornerstone for NPR-based anime style conversion. Instead of relying on pre-existing data, AnimeGANv2 employs a diverse collection of real-world images captured from various sources. This collection includes photographs spanning different scenes, environments, and subjects, providing a rich and comprehensive training corpus for the model.

The process of NPR in our project entails several intricate steps:

1) *Data Preprocessing:* Raw input images undergo preprocessing procedures to standardize their dimensions, color profiles, and quality. Techniques such as image normalization, resizing, and color correction are applied to ensure consistency and compatibility with the AnimeGANv2 model.

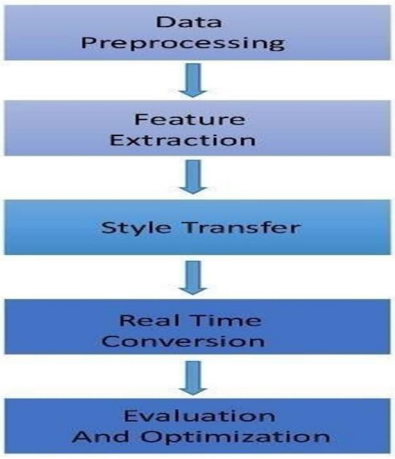


Fig .1. Procedure of NPR

2) *Feature Extraction:* AnimeGANv2 leverages powerful convolutional neural network (CNN) architectures to extract high- level features and representations from the input images. These features capture essential characteristics such as shapes, textures,colors, and spatial relationships, crucial for effective style transfer.
3) *Style Transfer:* Stylization algorithms embedded within AnimeGANv2 enable the emulation of distinct anime styles, including Hayao Style, Shinkai Style, and Paprika Style. Each style offers unique characteristics and aesthetics, ranging from vibrant colors and intricate details to bold outlines and surreal imagery. These algorithms manipulate the extracted features tomatch the aesthetic attributes and visual cues characteristic of each anime style, resulting in faithful and expressive transformations.
4) *Real-time Conversion:* The stylized features are seamlessly integrated back into the input images, facilitating real-time conversion of photographs into anime-style artworks. Through efficient processing pipelines and optimized algorithms, users can experience instant and immersive transformations, enhancing their creative workflow and artistic endeavors.
5) *Evaluation and Optimization***:** The generated anime-style images undergo rigorous evaluation and optimization processes to assess their visual fidelity, coherence, and artistic appeal. Quality metrics, user feedback, and subjective assessments guide iterative improvements and refinements, ensuring superior output quality and user satisfaction.

B. Neural Network-Based Stylization Techniques for Anime Artwork

Our project delves into the realm of Neural Network-based Stylization Techniques for Anime Artwork, aiming to revolutionizeanime-style conversion. Utilizing a diverse collection of over 10,000 real-world photographs and curated cartoon imagery, our model, AnimeGANv2, harnesses the power of deep learning for efficient and expressive style transfer. With from vibrant colors and intricate details to bold outlines and surreal imagery.

In Anime GANv2, we emulate a variety of iconic anime styles, including Hayao Style, Shinkai Style, and Paprika Style. These styles encapsulate the essence of renowned anime aesthetics, ranging Throughout the training process, AnimeGANv2 under goes iterative refinements, integrating cutting-edge techniques such as edge-promoting adversarial loss and semantic content loss functions.

©IJRASET: All Rights are Reserved



Fig .2. Different Styles of Anime Art-Work Hayao, Shinkai, Paprika

The model's architecture features 8 linear bottleneck residual blocks, ensuring efficient feature extraction and faithful style representation. In performance evaluation, we employ standard metrics such as the Structural Similarity Index (SSI) and Peak Signal-to-Noise Ratio (PSNR) to assess the quality of converted images. Additionally, user feedback supplements our evaluation process, guiding further enhancements and optimizations.

C. Image Synthesis Approaches Using AnimeGANv2

In our project, we harness the power of Generative Adversarial Networks (GANs) for image synthesis, specifically focusing on anime-style conversion. GANs have emerged as a potent tool for various image-related tasks, including text-to-image translation, image inpainting, and image super-resolution. The fundamental idea behind GANs involves training two neural networks simultaneously: a generator and a discriminator. The generator fabricates images, while the discriminator evaluatestheir authenticity. Through adversarial training, the generator learns to produce images that closely resemble real images, gradually improving its ability to generate high-quality outputs.

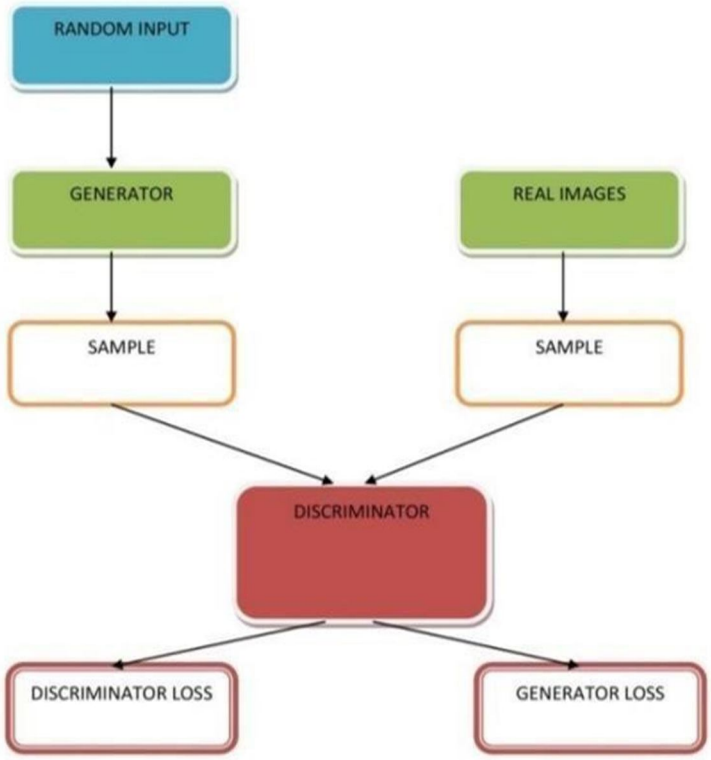


Fig .3. Flow Chart Showing How Generator and Discriminator Work in AnimeGanv2

D. AnimeGanv2 Structure

The architecture of AnimeGANv2 is designed to facilitate the transformation of real-world photographs into anime-style images. At its core, AnimeGANv2 employs a Generative Adversarial Network (GAN) framework, which consists of two primary components: the generator network (G) and the discriminator network (D).

The architecture of AnimeGANv2 in this project comprises a sophisticated generator network designed to transform real- world photographs into captivating anime-style images. The generator network consists of a series of convolutional layers and activation functions, strategically arranged to capture the intricate details and distinctive characteristics of anime artwork.

Initially, the input images undergo convolutional operations with a kernel size of 7x7, followed by max-pooling with a stride of1 and a channel size of 32. Subsequently, the features are further refined through additional convolutional layers with varying kernel sizes and channel sizes, including 3x3 kernels with a stride of 2 and 64 channels and 3x3 kernels with a stride of 1 and 128 channels. The network architecture also incorporates residual blocks, inspired by the layout proposed in previous studies, to enhance feature extraction and preserve essential details during the transformation process.

Finally, the output layer utilizes convolutional operations with 3x3 kernels to reconstruct the anime-style images with exquisite precision and fidelity. Similarly, the discriminator network in AnimeGANv2 comprises multiple convolutional layers to evaluate the authenticity of generated anime-style images. It starts with a 3x3 convolutional layer with a stride of 1, followed by a layer with a 3x3 kernel and a stride of 2. These initial layers capture low-level features and spatial information in real and synthetic anime images. Deeper layers feature 128 channels, 3x3 kernels, and 1-stride, enabling analysis of higher-level features and complex patterns in anime artwork. Additional convolutional layers include 256 channels and 3x3 kernels with 1stride.

The network concludes with a single-channel convolutional layer, outputting a scalar value representing image authenticity. AnimeGANv2's discriminator effectively discerns real anime images from synthetic ones, ensuring high-quality anime-style artwork generation. The discriminator network (D) in AnimeGANv2 is vital for determining the authenticity of input images as real anime images. It utilizes a patch-level discriminator, focusing on local features essential for distinguishing between real and synthetic anime images.

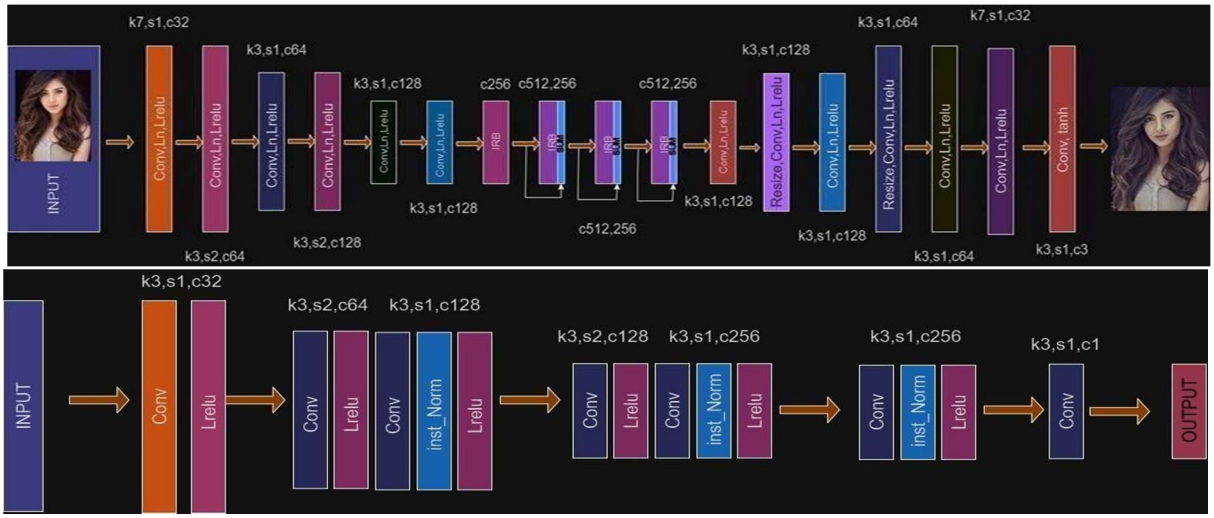


Fig .4. Architecture of Generator and Discriminator Networks in Proposed AnimeGanv2

The network architecture is intentionally kept shallow, beginning with flat layers and followed by two strided convolutional blocks to reduce resolution and capture crucial local features. A feature construction block and a 3x3 convolutional layer are then employed to derive the classification response.

E. Adversarial Loss

The adversarial loss function is a key component used to train the generator network (G) to generate anime-style images that are indistinguishable from real anime images. It works by pitting the generator against the discriminator network (D), which aims to differentiate between real anime images and those generated by G. Formally, the adversarial loss function is defined as:

$$f_{\text{av}}(G, D) = Ex{\sim}p\text{dt}(x)[\log D(x)] + Ez{\sim}(z)[\log(1- D(G(z)))]$$

G is the generator network. D is the discriminator network. x represents real anime images. z represents the noise vector. $p_{dt}(x)$ is the distribution of real anime images. p(z) is the distribution of the noise vector. The adversarial loss encourages the generator to produce images that are realistic enough to fool the discriminator, while the discriminator aims to correctly classify real and generated images. This adversarial training process leads to the generator learning to produce high-quality anime-style images.

F. Feature Matching Loss:

The feature matching loss is a component used to enhance the training of the generator network (G) by encouraging it to generate Images that not only fool the discriminator but also match the features of real anime images at multiple layers of abstraction. The feature matching loss is computed by comparing the features extracted by intermediate layers of the discriminator for bothreal and generated images. It measures the discrepancy between the features of real anime images and those of the generated images. Formally, the feature matching loss can be defined as follows:

$$\mathscr{L}_F(G, D) = \frac{1}{N} \sum_{i=1}^{N} \| (x) - Di(G(z)) \|_1$$

The feature matching loss encourages the generator to produce images that not only fool the discriminator but also match the features of real anime images at various levels of abstraction, resulting in more visually appealing and realistic output images.

IV. IMPLEMENTATION

1) *Real-World Image Collection*: Real-world images were collected from various sources, representing a diverse range of scenes, subjects, and compositions. These images serve as the input dataset for the image-to-anime conversion experiments.
2) *User Authentication and Authorization:* A user authentication system was integrated into the web page interface to verify user identities and enforce access controls. Only authenticated users with appropriate permissions were allowed to upload images and initiate the conversion process.

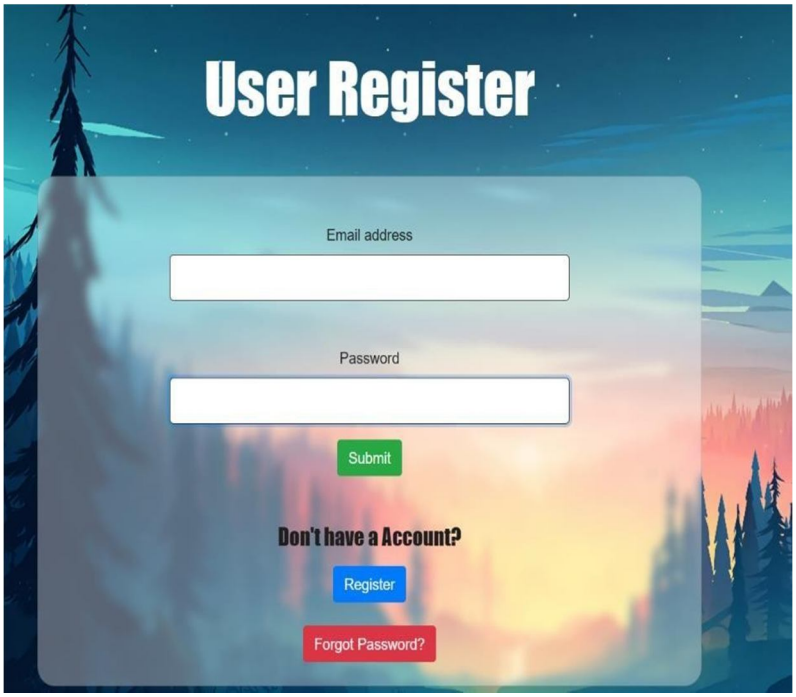


Fig .5. User Authentication System in Our Web Page

3) *Web-Based Image Upload:* A custom web page interface was developed to facilitate the uploading of real-world images for conversion to anime-style artwork. Users were able to securely upload images through the web page, which served as the interface for interacting with the image conversion model.

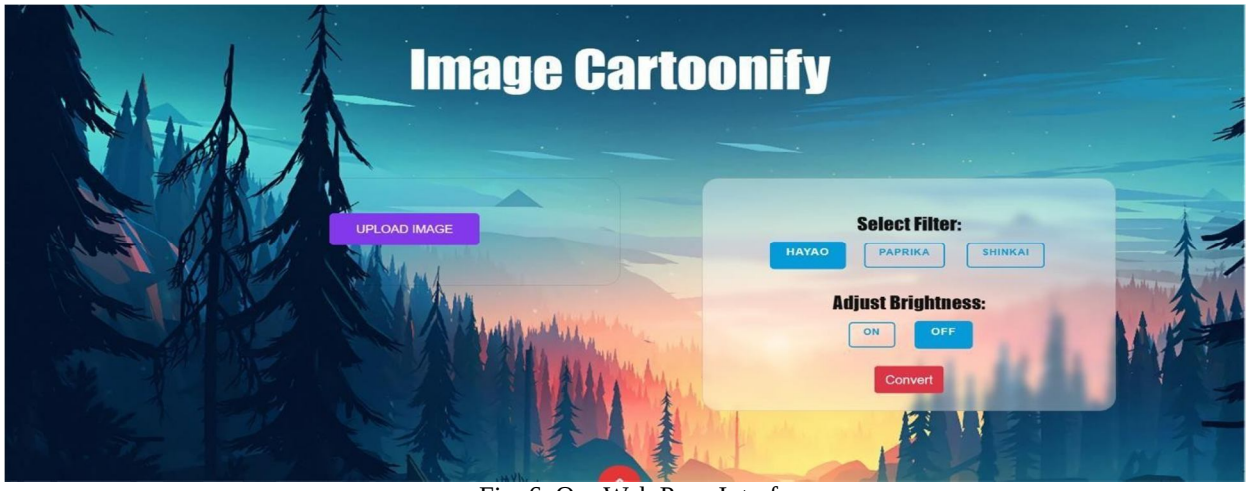


Fig .6. Our Web Page Interface

4) *Image-to-Anime Conversion: Web-Hosted Model:* The image-to-anime conversion model was deployed on a web server, allowing users to access the conversion functionality remotely through the web page interface. This approach enables seamless interaction with the model without requiring users to install or run the model locally.
5) *On-Demand Conversion:* Upon uploading an image through the web page, the image to anime conversion model processedthe input image in real-time, generating a corresponding anime-style rendition. Users could visualize the converted image directly within the web page interface.
6) *Secure Data Handling:* Data Privacy and Security: Stringent measures were implemented to ensure the privacy and securityof uploaded images. This included encryption protocols, secure data transmission channels, and access control mechanisms to safeguard user data and prevent unauthorized access.

V. RESULTS



Fig .7. Real Time Conversion in Our Web Page

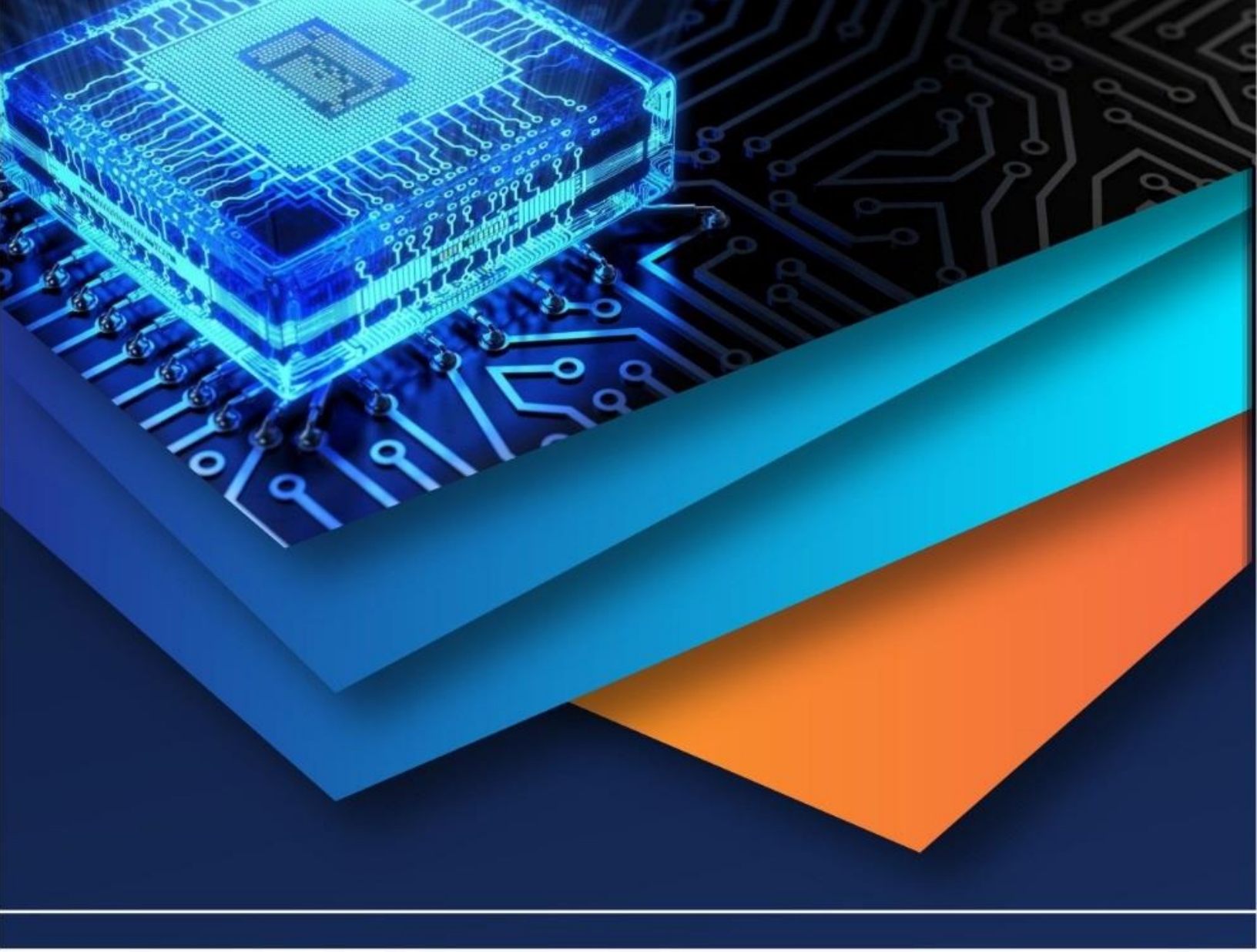
VI. CONCLUSION

In conclusion, AnimeGANv2 offers a cutting-edge solution for converting images into high-quality anime representations, drawing inspiration from iconic styles such as those by Makoto Shinkai, Miyazaki Hayao, and Paprika. With its advanced deep learning architecture, AnimeGANv2 empowers users to seamlessly transform ordinary photographs into captivating anime artworks. As the demand for high-quality anime-style imagery continues to rise, AnimeGANv2 stands at the forefront, promising to redefine digital artistry and inspire creativity in the realm of visual storytelling.

©IJRASET: All Rights are Reserved

REFERENCES

[1] Chao Dong, Chen Change Loy, and Xiaoou Tang," Accelerating the Super-Resolution Convolutional Neural Network," doi:10.1007/978-3-31946475-6_25, vl:9906, October 2016.
[2] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1132-1140, doi:10.1109/CVPRW.2017.151.
[3] Y. Chen, Y.-K. Lai, and Y.-J. Liu, "CartoonGAN: Generative Adversarial Networks for Photo Cartoonization," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, USA, 2018, pp.9465–9474, doi:10.1109/CVPR.2018.00986
[4] Y. Chen, Y.-K. Lai, and Y.-J. Liu, "Transforming photos to comics using convolutional neural networks," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 2010–2014, doi:10.1109/ICIP.2017.8296634.
[5] Yugang Chen, Muchun Chen, Chaoyue Song, and Bingbing Ni" CartoonRenderer: An Instance-based MultiStyle Cartoon Image Translator," November 2019.
[6] X. Zhao, Y. Zhou, J. Wu, Q. Xu, and Y. Zhang, "Turn Real People into Anime Cartoonization," 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2021, pp. 337–339, doi:10.110 9/ICCECE51280.2021.9342433.
[7] Shu et al., "GAN-Based Multi-Style Photo Cartoonization," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 10, pp. 3376- 3390, 1 Oct. 2022, doi:10.1109/TVCG.2021.3067201.
[8] J. Acharjee and S. Deb, "Cartoonize Images using TinyML Strategies with Transfer Learning," 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Zallaq, Bahrain, 2021, pp. 411-417, doi:10.1109/3ICT53449.2021.9581835.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)