Image Restoration and Enhancement Using Generative AI

Sindhuja Boda
Computer Science and Engineering
GRIET, Hyderabad, India
sindhuboda777@gmail.com

Phanindar
Computer Science and Engineering
GRIET, Hyderabad, India
phani0120@gmail.com

Tadepalli Siva Abhay Kiriti
Computer Science and Engineering
GRIET, Hyderabad, India
tskiriti21@gmail.com

Vadlakonda Sravan Kumar
Computer Science and Engineering
GRIET, Hyderabad, India
sravanv284@gmail.com

Surla Nagendra
Computer Science and Engineering
GRIET, Hyderabad, India
snani10323@gmail.com

***Abstract*— High-quality images are increasingly in demand in various sectors, mostly even in areas that focus on accuracy to the highest extent, such as healthcare and security; however, common image degradation modes include noise, blurring, and low resolutions that affect not only the resolution but also usability. The main disadvantage with conventional enhancement techniques is that they cannot satisfactorily recover most degraded images. The problem here is overcome using Generative Adversarial Networks (GANs) combined with the Deep Generative Prior (DGP) model, providing a solid solution for image restoration and enhancement based on DGP, enabling a relatively more flexible and adaptive approach than traditional GAN-based methods. This article introduces a unified framework combining three multidisciplinary deep learning approaches with respect to image processing: DeOldify for image colorization, ESRGAN for image enhancement, and GFP-GAN for restoration of facial images. Moreover, we shed light on how each component was implemented, the methodology used, and the outcomes acquired, revealing how all of them contribute together when combined in an integrated processing pipeline for degraded images. In this paper, GAN capabilities are improved by iteratively finetuning the generator for every target image, improving versatility and performance of various tasks such as super resolution, colorization, and inpainting. Our paper describes our methodology involving an integration of GAN inversion techniques, attention mechanisms, and residual learning for the production of high-quality, artifact-free images. This is a well-suited model for real-time applications with a viable solution for critical tasks requiring enhanced image quality. By leveraging these advancements, we demonstrate significant improvements in image fidelity and processing speed. Furthermore, our approach not only addresses the common challenges faced in image restoration but also sets a new benchmark for future research in the field.**

Keywords*— *Image Restoration, Enhancement, Generative Adversarial Networks (GANs), Deep Generative Prior (DGP), Super-Resolution, Colorization, Progressive Fine-Tuning.

I. INTRODUCTION

Recently, high-quality image restoration and enhancement have become essential across a wide range of applications, including healthcare, security, and environmental monitoring, where the clarity and detail of images are crucial. Nature and scenic images have special requirements to provide maximum fidelity in their textures, colours, and detailed information for an accurate representation of landscapes and natural scenes. However, such images are normally highly degraded and therefore of low value for automated analysis as well as human interpretation: Many lose too much information that is not noise-like, sometimes becoming quite blurry; they lose resolution; many suffer colour loss; and the above noisiness, caused by unnecessary loss of information, applies even to clean images. Some restoration approaches have exploited the potential of filtering techniques as well as the CNN, but these often fail to tackle all degradation types effectively due to their rigidity. These traditional methods often do not provide photorealistic, high-quality restorations, especially when there are complex and detailed scenes of images, which are typical for most nature scenes.

Recently, Goodfellow et al. presented a new approach in the era of GANs on image processing that has achieved promising success in super-resolution, colorization, and inpainting among other applications. GANs support the adversarial approach where a generator network builds up images, while a discriminator network tries to differentiate the generated images from the real ones. This adversarial architecture enables GANs to learn complex data distributions, and the outputs from such models can closely resemble actual images. For instance, SRGAN as well as many other forms of GANs can substantially improve images while showing promise for further developments. However, mode collapse, artifact creation, and bad adaptability to images with mixed, complex types of degradation affect these GAN-based methods. Such issues lead to unnatural texture images and loss of structural identity, which are especially problematic on nature or landscape images where both detail and reality are intended to be crucial.

In pursuit of these limitations, one recent research effort has focused on the proposal of Deep Generative Prior (DGP), which takes the lead from GAN inversion methods to introduce added flexibility in image restoration. DGP uses a pre-trained GAN as a general-purpose image prior that would capture the textures, colours, and so on, at a high level of semantics. In contrast to all previous methods of GAN inversion, which simply kept the generator fixed, DGP involves progressive fine-tuning where, for each specific target image, the generator parameters are adapted. The adaptive nature of this method is specifically useful for nature images while capturing fine-grained details, preserving natural textures, and minimizing unrealistic artifacts. Restoration remains within the natural image manifold, regarding discriminator-guided feature matching loss in DGP, making it visually more accurate and realistic.

This paper is aimed at exploring GANs with deep generative prior applied for the process of image restoration and enhancement of scenery along with natural photographs. In this progressive fine-tuning, attention mechanisms along with residual learning are combined to address the different requirements of the nuanced process of natural image processing. Under these considerations, the proposed model shall be capable enough of handling a wide variety of degradation types, thereby bringing about flexibility in improving image quality where detail and accuracy are a priority. In this paper, we will facilitate the establishment of GAN-based restoration methods for high-quality restorations that capture unique characteristics of natural scenes.

This paper presents three complementary deep learning-based approaches that together provide a comprehensive solution for image transformation and restoration.

1) DeOldify: A NoGAN-based colorization system that transforms grayscale images to colour by employing self-attention mechanisms and perceptual loss functions to maintain spatial coherence and generate natural-looking colours.

2) ESRGAN: A super-resolution generative adversarial network for enhancing image resolution and quality through Residual-in-Residual Dense Blocks (RRDB) and high-order degradation modelling to effectively handle real-world image degradations.

3) GFP-GAN: A facial restoration system that leverages generative priors from a pre-trained StyleGAN2 model to recover detailed facial features while preserving identity, utilizing Channel-Split Spatial Feature Transform (CS-SFT) for adaptive feature modulation.

By combining these specialized approaches, we aim to address the various aspects of image degradation simultaneously, providing a holistic solution for high-quality image restoration. The integration of these methods allows us to handle multiple degradation types—from color loss to resolution reduction and detail deterioration—in a complementary fashion, resulting in a comprehensive image enhancement preserving the spatial characteristics of the original scenes, consequently ameliorating their quality and usability.

With the dawn of technology concerning artificial intelligence and deep learning, the fortification of the real-time application seems to be set in the fields of autonomous driving, medical imaging, and digital media restoration. These very technologies contribute toward upholding historical records, augmenting security systems, and enhancing visual quality in images used in industries.

II. LITERATURE SURVEY

In [1], Deep the researchers explored the use of deep generative priors from GANs trained on large-scale natural images for versatile image restoration and manipulation tasks, such as colorization, inpainting, and super-resolution. The model fine-tunes the GAN generator progressively, allowing it to restore missing semantics and textures in a natural and faithful manner, achieving strong results on tasks involving complex images like those from ImageNet.

In [2], it introduces a novel method for restoring degraded images using multiple frames and Generative Adversarial Networks (GANs). The technique trains a generator to learn clean images from several degraded frames while a discriminator ensures image quality. This multi-frame approach enhances the restoration of images affected by complex degradation patterns like noise and blur. The method outperforms traditional techniques in terms of image quality and consistency, making it effective for applications such as video restoration and surveillance. The approach also maintains temporal coherence across frames.

In [3], authors attempted to address problems of multimodal ambiguity and color bleeding in image colorization. The authors have devised a framework called Pal GAN that incorporates palette estimation with chromatic attention for preserving color coherence and edge preservation at the time of color assignment to grayscale images. Results in the test beds of ImageNet and COCO-Stuff resulted in Pal GAN surpassing the previous models in FID scores as well as LPIPS scores while providing diverse, realistic, and controllable results.

In [4], the authors proposed an image quality enhancement framework that combines adaptive filtering with bidirectional memory and spatiotemporal constrained optimization. The approach adaptively adjusts filtering parameters according to local image content, preserving subtle details effectively. A bidirectional memory module utilizes past and future contextual information to steer the enhancement process, and spatiotemporal constrained optimization imposes consistency across spatial areas and temporal frames. Experimental results show that this method far surpasses conventional methods by eliminating artifacts and retaining fine details as well as global image coherence.

In [6], It uses a multi-stage framework, where images are restored in several steps, with each stage refining the output of the previous one. The model employs a U-Net-based architecture combined with advanced attention mechanisms for more effective feature extraction and noise removal. This approach allows for superior performance in restoring images with various types of degradation, including noise, blur, and low resolution, making it useful in practical applications like Photography and Medical Imagin.

In [7], the authors experimented with a new multi-scale residual architecture, dubbed MIRNet-v2, to improve state-of-the-art performance in deblurring, denoising, and super-resolution tasks in image restoration. MIRNet-v2 not only preserves high-resolution spatial details but also utilizes information from multi-scale contexts, establishing it as one of the best-performing models for six real-image benchmark datasets, excelling in both speed and accuracy.

In [8], the authors proposed a GAN-based model known as SRSRGAN that is specifically designed to perform both underwater image restoration and super-resolution simultaneously. This design accommodates a two-stage structure: the first stage focuses on noise removal and color degradation correction, while the subsequent super-resolution stage restores high-frequency details. Experiments on datasets such as UIEBD and UFO-120 showed that SRSRGAN outperforms state-of-the-art methods across both restoration and super-resolution tasks.

The paper [9] presents a deep, fully convolutional auto-encoder network for image restoration tasks like denoising, super-resolution, and inpainting. The network is designed with symmetric convolutional-deconvolutional layers linked by skip connections, which help preserve image details while removing corruptions. These skip connections accelerate training and improve restoration results. This method significantly enhances image quality, outperforming conventional techniques in multiple tasks.

In [10], the authors presented SwinIR, an image restoration model that utilizes the Swin Transformer model to solve tasks like super-resolution, denoising, and artifact removal. The approach follows a hierarchical transformer architecture that divides images into non-overlapping windows, enabling efficient local and global feature modeling by using shifted window mechanisms. Through the fusion of multi-scale feature extraction and deep transformer blocks and residual connections, SwinIR efficiently recovers fine image details and texture. Large-scale experiments on test datasets prove the proposed method beats conventional CNN-based approaches with a new state of the art on image restoration tasks.

III. METHODOLOGY

1) *Image Colorization Using DeOldify*

DeOldify is a deep-learning-based implementational framework that seeks to colorize grayscale images automatically. Artificial Intelligence-powered DeOldify uses a sophisticated architecture of generative deep learning neural networks, which converts a black-and-white image into its color presentation most likely. The architecture of DeOldify is not based on the traditional Generative Adversarial Networks (GANs) improvements. The implementation employs a "NoGAN" training methodology, which combines aspects of transfer learning and progressive training. The model architecture consists of:

->A generator network based on a modified U-Net architecture with ResNet backbone

->Self-attention mechanisms to enhance spatial coherence .

->Perceptual loss functions that leverage pre-trained networks to evaluate image similarity.

2) *Image Enhancement Using ESRGAN*

The Real-Enhanced Super-Resolution Generative Adversarial Network (RealESRGAN) represents a significant advancement in the field of image enhancement. This implementation utilizes deep learning techniques to effectively upscale low-resolution images while simultaneously reducing noise and improving overall visual quality. The model builds upon the foundational ESRGAN architecture with specific modifications to enhance performance on real-world degraded images.The RealESRGAN operates on the framework of adversarial learning comprising two main networks:

1. <u>TheGeneratorNetwork</u>: a deep residual network with Residual-in-Residual Dense Blocks (RRDB) and responsible for the actual super-resolution task.
2. <u>TheDiscriminatorNetwork:</u> A VGG-style convolutional network that distinguishes between high-resolution images generated and real high-resolution images.

The generator uses various network techniques:- Improved residual connections allowing gradient information to flow more easily.- Dense feature fusion to make the most of the available information.- Sub-pixel convolution layers allowing upsampling to be performed very efficiently.

3) *Image Restoration Using GFP-GAN*

GFP-GAN indicates a stout step forward into blind face restoration, which combines beforehand verified facial generative priors packed into the GAN architecture for the filling of very severely degraded facial images with good fidelity. It intends to fix the arduous problems about the high-frequency detail recovery and identity consistency in face restoration tasks.

GFP-GAN adopts an innovative architecture consisting of two key components:

-><u>Generative facial prior</u>-a pre-trained StyleGAN2 model that encompasses rich knowledge about the prior of faces

☾ <u>U-shaped generator</u>, which is a feature transformation network that connects degraded image features with the latent space of the generative prior.

The architecture further incorporates adaptive feature modulation by means of the channels-SFT layers, facial component enhancer modules for the preservation of details, and multi-scale identity maintenance mechanisms to ensure facial identity integrity.

IV. IMPLEMENTATION

The implementation consists of a structured pipeline involving a combination of three deep-learning architectures—DeOldify (colorization), RealESRGAN (super-resolution), and GFP-GAN (face restoration)—for comprehensive image enhancement. A stepwise description is given below:

A. *Environment setup: -*

Install the necessary dependencies, including PyTorch, OpenCV, and FastAI. Clone the repositories for DeOldify, RealESRGAN, and GFP-GAN. Ensure that the system has a compatible GPU for optimal performance.

B. *Model initialization and acquisition of pre-trained weights: -*

Load the DeOldify model for grayscale image colorization. Initialize RealESRGAN with a 4x upscaling factor for image super-resolution. Set up GFP-GAN to restore facial details using a pre-trained StyleGAN2-based facial prior.

C. Image Processing Pipeline: -

Image enhancement begins with colorization using DeOldify, which accepts a grayscale image into the model, and with a render factor setting that would control the vividness and detail of the colorization that comes up as the output. Once the secant is optimized, an output of the enhanced, colorized image is yielded and saved. The super-resolution process uses the RealESRGAN to apply an advanced upscaling model to low-resolution images in order to improve crispness while removing noise and artifacts, yielding a high-resolution output. In the face restoration step, GFP-GAN aims to emphasize facial details in degraded or very old photographs. Mainly, by delivering the model a distorted image and giving a go-ahead based on a pre-trained StyleGAN2 about a facial prior, it can proceed to fix the facial features while taking identity into account, rendering a polished, natural-looking output with improved sharpness and textures.

D. Output visualization and evaluation: -

Display the original image and the processed images for comparison. Colour accuracy, sharpness, and identity preservation will be evaluated in the output. Thus, qualitative and quantitative analysis to assess model performance.

E. Optimization and fine-tuning: -

Hyperparameters such as render factor (DeOldify), scale factor (Real-ESRGAN), and identity preservation level (GFP-GAN) will be modified. Experience different pre-trained weights to get finer results, thereby maximizing speed and efficiency optimization within the processing pipeline in applications on a large scale.

V. ARCHITECTURE

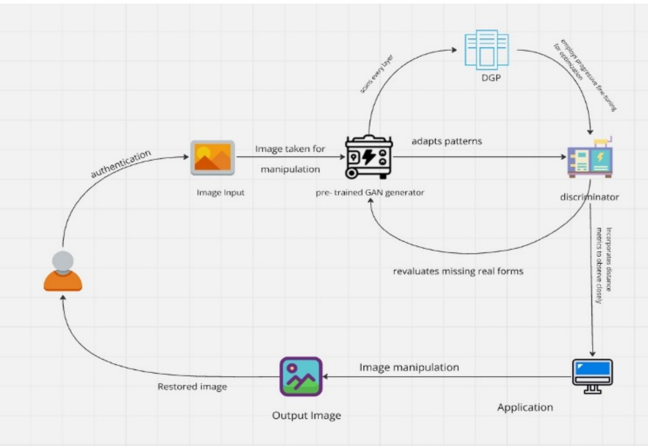

Fig. 1. Architecture

The architecture of the three-image generation features of Image Colorization, Image Enhancement, and Image Restoration is represented by a diagram. It shows a procedure wherein the input image is processed using a pre-trained generator of GAN, which modifies-presumably symbolizes-a deep generative model. The model itself carries out the item of a `discriminator'-to estimate and refine forms missing from the realism point of view. It thus manipulates the image to further enhance-details. The output is then delivered in the form of the final restored image.

VI. RESULTS AND DISCUSSIONS

In this experiment, it was confirmed that the grayscale image had been converted into a color image having near-exact natural-skin tones. The algorithm has genuinely restored the blue sky, green grass, and spotted pattern of the cheetah to look appealing. Close attention is attention-oriented regarding woolly texture details and picturesque backgrounds, maintaining their structural integrity. Results validate the effectiveness of the proposed model in producing high-fidelity colorization while preserving artifacts to a lesser extent.

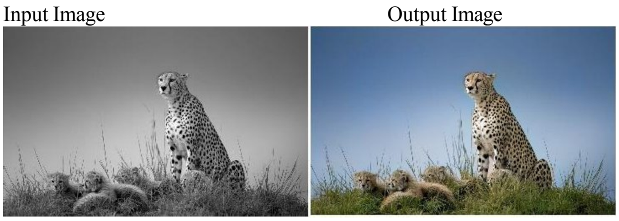

Fig. 2. Image Colorization

In short, the reconstructed image successfully conserves the missing or damaged portions and other associated features and texture qualities of the face. This restoration has greatly improved the clarity of the image, reducing such apparent features as cracks and scratches while enhancing a certain level of detail. Evidently, certain features of the face-such as eyes, lips, and hair-have been smoothed and made to look natural. There are still some minor artifacts present in those heavily degraded areas, indicating areas that need some more improvement. The model performs great in maintaining structural consistency along with improving perceptual quality.

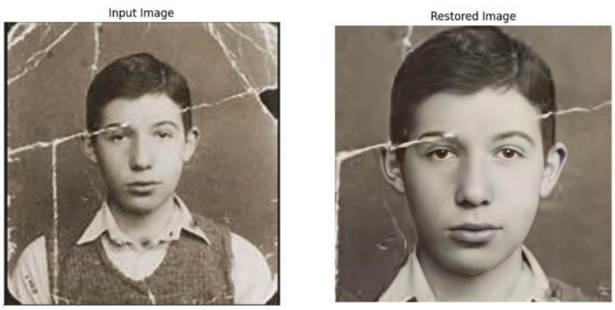

Fig. 3. Image Restoration

The comparison between the noisy and denoised images shows a significant reduction in noise while maintaining the structural integrity of the original scene. The denoised image exhibits enhanced brightness, sharper edges, and improved color contrast. The details in the sky, mountains, and vegetation appear smoother and more visually appealing. Moreover, the suppression of noise does not introduce noticeable artifacts, indicating the robustness of the enhancement approach.

Fig. 4. Image Enhancement

VII. CONCLUSION

Techniques of image processing such as image colorization, image restoration, and image enhancement help improve the quality of an image and thus make it more useful in various fields.

->Image Colorization is a method of applying actual colors to black-and-white photographs so that old pictures appear new and upgraded with many more details. It is useful for the restoration of historical images, medical scans, and the enhancement of old films.

->Image Restoration is an art of restoring images that are damaged and blurry through scratches, noise, and distortion. It sees wide applications in the restoration of old photographs, forensic investigations, and medical imaging.

->Image Enhancement is very important because it makes the image clearer by enhancing brightness, contrast, and sharpness. It is very much used in satellite imaging, security cameras, and medical diagnostics.

VIII. REFERENCES

[1] Pan, X., Zhan, X., Dai, B., Lin, D., Loy, C.C. and Luo, P., 2021. Exploiting deep generative prior for versatile image restoration and manipulation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(11), pp.7474-7489.

[2] Velammal M., Singh T.I., Patil N.M. and Pal S. 2023. Multiframe image restoration using generative adversarial networks. *ICTACT Journal on Image and Video Processing*, 14(1), pp.3043-3048.

[3] Wang, Y., Xia, M., Qi, L., Shao, J. and Qiao, Y., 2022, October. PalGAN: Image colorization with palette generative adversarial networks. In European Conference on Computer Vision (pp. 271-288). Cham: Springer Nature Switzerland.

[4] Soundararajan S., Pushpalatha M., Charan P., and Nishant N. 2023. Improving image quality through adaptive filtering enhancement using bidirectional memory and spatiotemporal constrained optimization. *ICTACT Journal on Image and Video Processing*, 14(1), pp.3035-3042.

[5] Xu, B., Zhou, D. and Li, W., 2022. Image enhancement algorithm based on GAN neural network. IEEE Access, 10, pp.36766-36777.

[6] Fei, B., Lyu, Z., Pan, L., Zhang, J., Yang, W., Luo, T., Zhang, B. and Dai, B., 2023. Generative diffusion prior for unified image restoration and enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9935-9946).

[7] Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H. and Shao, L., 2022. Learning enriched features for fast image restoration and enhancement. IEEE transactions on pattern analysis and machine intelligence, 45(2), pp.1934-1948.

[8] Wang, H., Zhong, G., Sun, J., Chen, Y., Zhao, Y., Li, S. and Wang, D., 2023. Simultaneous restoration and super-resolution GAN for underwater image enhancement. *Frontiers in Marine Science*, *10*, p.1162295.

[9] Mao, X.-J., Shen, C., & Yang, Y.-B., 2016. Image restoration using convolutional auto-encoders with symmetric skip connections.

[10] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L. and Timofte, R., 2021. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1833-1844).

[11] Christopher, A., Hari Kishan, R., & Sudeep, P.V. (2022). Machine Learning Algorithms for Signal and Image Processing. In D. Ghai, S. L. Tripathi, S. Saxena, M. Chanda, & M. Alazab (Eds.), Machine Learning and Deep Learning for Emerging Trends in AI: Applications for Healthcare, Smart Environment, and Robotics (First ed., Ch. 5). Wiley.

[12] Zeng, D., Jiang, W., Yan, X., Fu, W., Shen, Q., Veldhuis, R., & Tang, B. (2024). Face super resolution with a high frequency highway. IET Image Processing.

[13] Das, S., Sen, S., Joardar, D., & Chakraborty, G. (2024). Image enhancement techniques to modify an image with machine learning application. In A. Dey, S. Nayak, R. Kumar, & S. N. Mohanty (Eds.), Emerging trends in AI and machine learning (Chapter 11). Wiley

[14] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L. and Timofte, R., 2021. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1833-1844).

[15] Christopher, A., Hari Kishan, R., & Sudeep, P.V. (2022). Machine Learning Algorithms for Signal and Image Processing. In D. Ghai, S. L. Tripathi, S. Saxena, M. Chanda, & M. Alazab (Eds.), Machine Learning and Deep Learning for Emerging Trends in AI: Applications for Healthcare, Smart Environment, and Robotics (First ed., Ch. 5).

[16] Aakerberg, A., Nasrollahi, K., & Moeslund, T. B. (2021). Real-world super-resolution of face-images from surveillance cameras. IET Image Processing, 16(2), 335-343.

[17] Wang, L., Li, J., Yokoya, N., Timofte, R. & Guo, Y. (2024) 'Guest editorial: Advanced image restoration and enhancement in the wild', IET Computer Vision, 18(4), pp. 435-438.

[18] Momen Tayefeh, M., Momen Tayefeh, M. & Ghahramani, S.A. (2024) Image inpainting for corrupted images by using the semi-super resolution GAN. arXiv.