International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-13 Issue-2, January 2024

A Joint Framework of GFP-GAN and Real-ESRGAN for Real-World Image Restoration

Mousumi Hasan, Nusrat Jahan Nishat, Tanjina Rahman, Mujiba Shaima, Quazi Saad ul Mosaher, Mohd. Eftay Khyrul Alam

***Abstract**: In the current era of digitalization, the restoration of old photos holds profound significance as it allows us to preserve and revive cherished memories. However, the limitations imposed by various websites offering photo restoration services prompted our research endeavor in the field of image restoration. Our motive originated from the personal desire to restore old photos, which often face constraints and restrictions on existing platforms. As individuals, we often encounter old and faded photographs that require restoration to revive the emotions and moments captured within them. The limits of existing photo restoration services prompted us to conduct this research, with the ultimate goal of contributing to the field of image restoration. To address this issue, we propose a joint framework that combines the Real-ESRGAN and GFP-GAN methods. Our recommended joint structure has been thoroughly tested on a broad range of severely degraded image datasets, and it has shown its efficiency in preserving fine details, recovering colors, and reducing artifacts. The research not only addresses the personal motive for restoring old photos but also has wider applications in preserving memories, cultural artifacts, and historical records through an effective and adaptable solution. Our deep learning-based approach, which leverages the synergistic capabilities of Real-ESRGAN and GFP-GAN, holds immense potential for revitalizing images that have suffered from severe degradation. This proposed framework opens up new avenues for restoring the visual integrity of invaluable historical images, thereby preserving precious memories for generations to come.*

***Keywords:** GFP-GAN, Real-ESRGAN. Deep learning, Visual Integrity, Adaptable Solution.*

Manuscript received on 29 December 2023 | Revised Manuscript received on 09 January 2024 | Manuscript Accepted on 15 January 2024 | Manuscript published on 30 January 2024.

**Correspondence Author(s)*

Mousumi Hasan*, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: mousumi.cse@baiust.ac.bd, ORCID ID: 0009-0006-7499-7669

Nusrat Jahan Nishat, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: jnishat93@gmail.com

Tanjina Rahman, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: rahmantanjina983@gmail.com

Mujiba Shaima, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: goodmorn_simi@yahoo.com

Quazi Saad ul Mosaher, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: saad.quazi44@gmail.com

Mohd. Eftay Khyrul Alam, Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Academic Block, Cumilla, Chittagong, Bangladesh. E-mail: eftay@baiust.edu.bd

I. INTRODUCTION

In the realm of digital image processing, "image restoration" is the process of enhancing degraded images by eliminating or reducing distortions like noise, blurring, and compression artifacts. This critical practice significantly contributes to improving visual communication and information transmission, with the fundamental aim of accurately recovering the initial, unaltered image. Various algorithms and techniques are employed to approximate and revive the genuine fundamental data contained in deteriorated images. The restoration process involves a comprehensive analysis of the degradation model, explaining the extent of image distortion, followed by the application of inverse operations to counter or rectify these distortions. An exact and systematic approach is essential for optimal results. Image restoration finds applications in historical preservation, artistic restoration, photography, medical imaging, surveillance, satellite imaging, and forensic analysis, playing a vital role in improving the quality and genuineness of images for more thorough comprehension, assessment, and depiction of visual information.

In recent years, image restoration has gained significant attention in computer vision, aiming to enrich the quality and intricate details of images degraded by factors such as noise, blurriness, compression artifacts, and low resolution. Deep learning-based approaches like Real-ESRGAN [1] and GFP-GAN [2] have individually demonstrated remarkable success in addressing image restoration tasks. However, challenges posed by real-world images with complex degradation patterns persist, necessitating further exploration and investigation. This article introduces a groundbreaking integrated structure that leverages the capabilities of Real-ESRGAN and GFP-GAN to address the challenge of restoring images in real-world scenarios. Real-ESRGAN, a top-of-the-line algorithm for single-image super-resolution, has shown outstanding results in restoring high-definition details. Meanwhile, GFP-GAN, a recent breakthrough in adversarial generative networks, has exhibited superior performance in restoring images by eliminating noise, blurring, and artifacts. The integration of these two models has the potential to revolutionize the field of image restoration by providing a comprehensive solution for real-world scenarios. The proposed collaborative framework utilizes the Generative Facial Prior (GFP) technique for blind face restoration in real-world scenarios, leveraging pre-trained face Generative Adversarial Network (GAN) models like Style-GAN [3], [4].

Despite the remarkable degree of variability provided by face GANs in generating authentic faces, integrating such generative priors into the restoration process remains a daunting task. Previous attempts involving GAN inversion [5], [6], [7] often resulted in low fidelity due to inadequate low-dimensional latent codes guiding accurate restoration.

To overcome these challenges, we suggest GFP-GAN with careful designs, achieving a decent balance of realism and fidelity in a single forward pass. Specific components of GFPGAN include a degradation removal module and a pre-trained face GAN serving as a facial prior, connected by direct latent code mapping and coarse-to-fine Channel-Split Spatial Feature Transform (CS-SFT) layers. The suggested CS-SFT layers successfully incorporate generative prior while retaining high fidelity, performing spatial modulation on a split of the features and allowing the left features to flow through directly for greater information preservation. Additionally, we apply the identity-preserving loss to boost fidelity and introduce face component loss with local discriminators to enhance perceptual facial details.

The Real-ESRGAN model, proficient in single-image super-resolution, employs an intricate generator network, combining convolutional layers and residual-in-residual dense blocks (RRDB) to enhance image resolution. The model integrates advanced methodologies like upsampling to manage diverse scale factors and enhance the preservation of spatial information. Real-ESRGAN's proficiency in image restoration holds great potential for future advancements in deep learning.

The proposed collaborative framework anticipates providing a more resilient and potent avenue to tackle the challenges of real-world image restoration. By combining the strengths of the two methodologies, the aim is to enhance image resolution, reduce noise, eradicate blur, and recover lost intricacies, delivering a comprehensive solution for real-world image restoration complications. The integration of these two methodologies presents a promising prospect for advancing the field of image restoration and achieving significant breakthroughs.

Through an exhaustive examination of the architectural design, instructional approach, and enhancement techniques employed in the collaborative structure, this manuscript offers a significant and impressive addition to the field of image restoration. The recommended method's effectiveness and superiority have been authenticated by exhaustive experimentation on different real-world image datasets, unequivocally indicating its ability to achieve remarkable restoration results. As such, the framework holds great promise for advancing image restoration techniques.

II. LITERATURE REVIEW AND BACKGROUND STUDY

Super-resolution [8], [9] [10], [11], de-noising [12] [41][42], [13], de-blurring [14], and compression removal [15] are all common components of image restoration. To generate visually appealing outcomes, the generative adversarial network is commonly used as loss supervision to push solutions closer to the natural manifold [16], [17], [18], but our work seeks to use pre-trained face GANs to generate facial priors (GFP). Restoring the face. Two common face-specific priors: geometry priors and reference priors are used to further enhance performance and are based on generic face hallucination [19], [20]. Facial landmarks [21], [22], face parsing maps [23], [21], and facial component heat maps [24] are some of the geometric priors. However, 1) such priors always deteriorate in real-world circumstances and need predictions from low-quality inputs. 2) They may not have sufficient information for restoration since they primarily concentrate on geometrical limitations. Instead, our adopted GFP does not use explicit geometry estimation from degraded images and contains enough textures within its pre-trained network. Typically, reference priors [25], and [26] use references of the same identity. DFDNet [27] offers to create a face dictionary of each component (such as the eyes and mouth) using CNN features to help with the restoration to get over this problem. While our GFP-GAN could reconstruct faces as a whole, DFDNet primarily concentrates on vocabulary components and declines in areas outside of its dictionary coverage (such as hair, ears, and facial contour). Furthermore, the GFP might supply rich and various priors encompassing geometry, textures, and colors, but the restricted capacity of the dictionary limits its variety and richness. To discover the nearest latent codes given an input picture, GAN inversion [28], [6]. [29], [5] previously took use of the generative priors of pre-trained GANs [3], [4]. The hidden coding of Style GAN [3] is repeatedly optimized using PULSE [7] until the separation between the outputs and inputs is below a certain threshold. To enhance the quality of the reconstruction, mGAN prior [5] makes repeated optimization efforts. However, since the low-dimension latent codes are inadequate to serve as a restoration guide, these techniques often result in pictures with poor quality. In contrast, the CSSFT modulation layers we propose allow for high fidelity by allowing previous inclusion on multiple resolution spatial information. Additionally, our GFP-GAN does not need costly repeated optimization during inference. Typically, channel split operation is investigated to create compact models and enhance model representation capabilities. While GhostNet [30] divides the convolutional layer into two pieces and makes use of fewer filters to produce intrinsic feature maps, MobileNet [31] suggests depth-wise convolutions. In DPN [32], the dual path design permits feature reuse and fresh feature discovery for each route, boosting its capacity for representation. Super-resolution also makes use of a related concept [33]. Although our CS-SFT layers have diverse operations and goals, they have comparable spirits. To achieve a suitable mix of realness and fidelity, we implement spatial feature transform [34], [35], [43] on one split and leave the left split as identity.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

III. METHODOLOGY

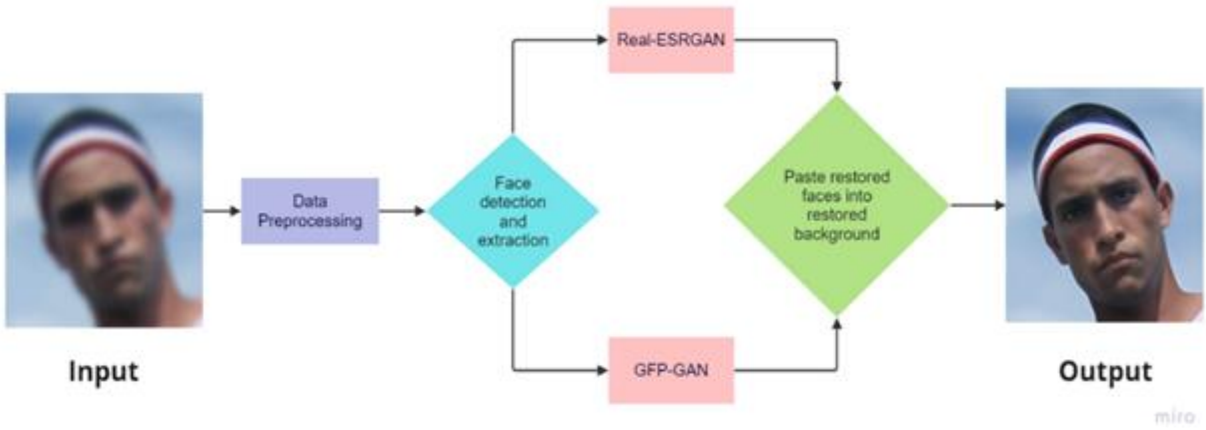

Fig 1: Overview of Proposed Methodology Framework

A. Data Preprocessing

Data preprocessing plays a pivotal role in optimizing the effectiveness of the training and testing processes. This critical stage involves various operations to generate low-quality images dynamically. The preprocessing pipeline includes random horizontal flipping, image degradation, random color jittering, conversion, and normalization.

- Random Horizontal Flipping: This operation is applied to ground truth (GT) images, introducing diversity to datasets and enhancing generalization. The process generates mirror images, exposing the model to distinct object orientations and modifications. Such an approach proves particularly valuable in training models for handling complex and diverse datasets.
- Image Degradation: GT images undergo image degradation to replicate practical image quality limitations. This involves blurring, downsampling, corruption with noise, and compression into JPEG format. Diverse degrees of blurring, downsampling, noise, and compression artifacts are introduced, enabling the model to develop robustness against these degradations. The dataset's versatility ensures the model's ability to comprehend and respond to a wide range of image degradations.
- Random Color Jittering: This technique introduces additional color variations to low-quality (LQ) images by randomly modifying their RGB values. The randomness enhances color diversity, allowing the model to handle various color distributions. Implementing this technique improves the model's capacity to generalize across different color representations, thereby enhancing overall performance.
- Conversion and Normalization: To ensure uniformity in data representation and facilitate efficient training, conversion and normalization steps are executed. Initially, images are converted from the BGR color space to the RGB color space, aligning with standard image representations for consistent data representation. Subsequently, the Height-Width-Channel (HWC) format is converted to the Channel-Height-Width (CHW) format, suitable for deep learning frameworks and efficient training. Finally, images are normalized by subtracting mean values and dividing by the standard deviation. This normalization scales inputs appropriately, enabling effective model training.

In summary, data preprocessing involves a comprehensive pipeline of operations that prepare images for training and testing. Each step contributes to the model's ability to generalize, handle diverse datasets, and exhibit robustness against various image degradations, ultimately enhancing overall performance and efficiency.

B. Overview of GFP GAN

The Generative Face Prior GAN (GFP-GAN) framework is employed for blind face restoration, aiming to restore a degraded facial image ***m*** without prior knowledge of the degradation type or degree. The primary goal is to estimate the highest-quality image ***n'***, closely resembling the ground-truth image ***n*** in terms of both realism and fidelity.

Figure 3.2 illustrates the comprehensive GFP-GAN framework. CS-SFT layers are used to integrate latent and spatial features, contributing to the generation of a revitalized image. The entire framework undergoes end-to-end training facilitated by a perceptual loss function, ensuring that the resulting image n'' is visually appealing and perceptually similar to the ground-truth image n. During inference, a degraded facial image m undergoes the degradation removal module, extracting latent and spatial features. These features then modulate the StyleGAN2 features [4], ultimately generating the revitalized image.

Following this, the process involves mapping ***Flatent*** to intermediate latent codes W through multiple linear layers. StyleGAN2, when presented with the latent code near the input image m, fabricates intermediate convolutional features denoted as FGAN. These features play a crucial role in providing rich intricate facial details captured within the pre-trained GAN weights. To achieve realistic and high-fidelity results, CS-SFT layers spatially modulate the face GAN features FGAN using multi-resolution features ***Fspatial*** in a coarse-to-fine approach.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In the training process, the introduction of facial component loss, along with discriminators, enhances perceptually significant facial components like the eyes, nose, and mouth. This approach contributes to the overall effectiveness of the GFP-GAN framework in blind face restoration.

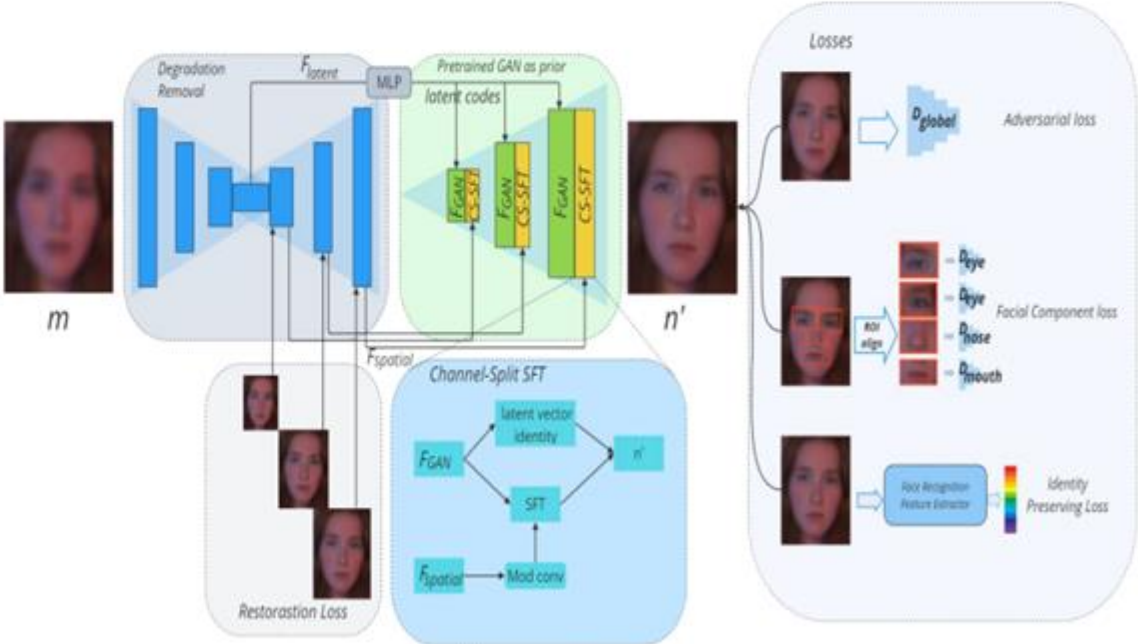

Fig 2: Overview of GFP GAN Framework

C. Degradation Removal Method

Restoring blind faces in real-life scenarios poses significant challenges due to intricate and severe degradation, encompassing color jitter, blur, grayscale conversion, down sampling, noise, and JPEG compression. Color jitter involves random variations in color values, often attributed to diverse lighting or camera configurations. The phenomenon of blur may arise from a shaky camera, motion blur, or the use of inferior-quality lenses. Converting an image to grayscale can lead to a loss of intricacy and surface quality. Down sampling results in decreased image resolution, potentially causing the loss of vital details, in contrast to upscaling, which enhances resolution. Additionally, noise may be introduced under dim illumination conditions, manifesting as roughness or dots in the image. The employment of JPEG compression for image compression is acknowledged to induce a loss of detail and artifacts, especially at elevated compression levels, making it a questionable choice for preserving delicate image details. These factors bear considerable significance in image processing, as they can profoundly influence the overall quality of the final output. Consequently, it is imperative to consider these factors when working with images to achieve optimal outcomes.

$$F_{latent}, F_{spatial} = U - Net(m) \quad (1)$$

To address the mentioned degradation and recover uncorrupted characteristics from the input image m, the U-Net architecture is employed. This includes latent features (Flatent) and multi-resolution spatial features (Fspatial). The technology's advantages are twofold: it expands the receptive field, capturing more information from input image m to eliminate significant blurs, and it forms multi-resolution attributes for detecting intricate details at various scales. This transformative technology significantly impacts image processing. Flatent maps the input image m to the nearest latent code in StyleGAN2, while Fspatial modulates StyleGAN2 features. To mitigate degradation, an L1 restoration loss is applied at each resolution scale during initial training. Output images for each U-Net decoder resolution scale are generated, constrained in proximity to the ground-truth image n's pyramid. These strategies contribute to enhancing the system's overall effectiveness and accuracy.

D. Generative Facial Prior and Latent Code Mapping

The utilization of pre-trained face GANs (Generative Adversarial Networks) is widespread for providing diverse and rich facial features tailored to a specific task. In this context, the pre-trained face GAN encapsulates a distribution over faces via learned convolution weights, constituting a generative prior. Consequently, the GAN incorporates information and structures of human faces, enabling the generation of innovative facial features resembling those observed in the training data. The conventional approach to leveraging generative priors entails mapping the input image m to its nearest latent codes Z, compressed representations of m, and subsequently generating the corresponding output through a pre-trained GAN [28], [6] [29], [5].

Pre-trained face GANs are a valuable resource for generating diverse facial features tailored to specific tasks, suggesting an alternative to directly producing the final image. Instead, it is recommended to generate intermediate convolutional features denoted as FGAN, capturing a higher degree of facial details. This intermediate step enhances fidelity in the final image n'. Utilizing the encoded vector Flatent of the input image m, produced by the U-Net as described in Equation 3.1, involves a multi-step process to better preserve semantic properties.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Specifically, the vector is initially mapped to intermediate latent codes W using a transformation from Z, accomplished with several multi-layer perceptron layers (MLP) [29]. These intermediate latent codes (We) are then processed through each convolution layer of the pre-trained GAN, generating GAN features for different resolution scales. It is noteworthy that the style code Z undergoes normalization before being fed into the MLP layers.

$$W = MLP(Flatent) \tag{2}$$

$$FGAN = StyleGAN(W) \tag{3}$$

By employing this approach, we can augment our preservation of the image's semantic characteristics, crucial for image restoration and enhancement. The utilization of intermediate latent codes ensures that the image maintains its original characteristics and features, preventing excessive distortion or modification. This methodology facilitates more precise and efficient image restoration and enhancement, resulting in improved outcomes and overall satisfaction. Harnessing the properties of Generative Adversarial Networks (GANs), the presented approach ensures a heightened level of intricacy obtained from the nearest face. The GAN features undergo modulation with input features, enhancing fidelity in the output image n'. This modulation process combines rich facial details from the GANs with pertinent information from the input image m, producing an output image n' that not only mirrors the input face but also exhibits enhanced visual characteristics. The proposed approach provides a time-efficient alternative to traditional methods by leveraging pre-trained GANs and generating intermediate convolutional features. Through this combination of techniques, results can be achieved considerably more swiftly than with traditional methodologies while preserving the quality of the ultimate outcome. Generative models possess the capacity to encapsulate a diverse range of elaborate priors extending beyond realistic details and vivid textures. These models encompass not only color priors but also offer the potential for face restoration and color enhancement. In real-world face images, such as old images, colors often manifest as black-and-white, vintage yellow, or dim. However, the dynamic color prior in generative facial prior allows for color enhancement, including colorization. Generative models contribute to the restoration and manipulation of face images by incorporating relevant information about the geometry and structure of the face.

E. Channel-Split Spatial Feature Transform

We leverage the spatial characteristics generated by the U-Net to modulate GAN features, thereby enhancing image quality. In the realm of face restoration, preserving spatial information from the inputs is paramount, as it plays a pivotal role in maintaining local characteristics for fidelity preservation and adaptive restoration across different spatial locations of a face. To accomplish this, we employ the Spatial Feature Transform (SFT) [34], a method that generates affine transformation parameters for spatial-wise feature modulation. SFT's primary objective is to compute inverse affine transformations for each face in the input image m, ensuring that the restored faces can be precisely aligned and pasted back onto the original image. SFT has demonstrated remarkable efficiency in integrating various conditions in image generation and restoration, underscoring its significance in the field of image processing. The application of input spatial characteristics from the U-Net with SFT has yielded substantial improvements in the quality and precision of the restored images, establishing it as an essential tool for both image restoration [34] and generation [35]. SFT serves as a technique for generating affine transformation parameters from input features *Fspatial* at every resolution scale. To derive (α, β), the affine transformation parameters and modulated convolution layers are utilized instead of directly obtaining them from input features *Fspatial*, as typically done in SFT. The modulation network is trained to generate (α, β) for each feature map in FGAN, taking *Fspatial* as input and outputting the desired transformation parameters. The equation for modulated convolution can be formulated as:

$$F_{output} = \alpha * (F_{GAN} \otimes w) + \beta \tag{4}$$

The standard convolution operation ⊗ is denoted, where w represents the learnable convolution filter. Subsequently, the modulated convolution layer employs (α, β) for scaling and shifting operations to modulate FGAN. This sophisticated approach distinguishes itself from previous methods, showcasing its effectiveness in spatial feature transformation. Through the modulation of the transformed feature map with the parameters (α, β), the modulated convolution operation seamlessly integrates spatial feature transformation. The layers of CS-SFT conduct spatial modulation on a segment of GAN features, specifically utilizing the input features *Fspatial* crucial for ensuring fidelity. Simultaneously, the remaining GAN features responsible for realness directly pass through without modulation, as illustrated in Figure 2.The output of the CS-SFT layers is denoted as F_{output} and is calculated using Equation 4.

$$F_{output} = CS\text{-}SFT(F_{GAN}|\alpha, \beta) \tag{5}$$

The CS-SFT technique, as described in the prior work [30], strategically integrates prior knowledge with input image modulation. This integration aims to strike a balance between texture fidelity and faithfulness, culminating in a realistically restored image that faithfully preserves the essential details of the original face. An additional advantage of CS-SFT is its ability to reduce complexity by necessitating fewer channels for modulation, akin to the GhostNet method [30]. In our restoration process, channel-split SFT layers were employed at each resolution scale, proving highly effective in attaining our restoration objectives.

F. Model Objective

The core objectives of our GFP-GAN training encompass the achievement of the following learning goals:

- Reconstruction Loss: This involves a combination of two distinct components—L1 loss and perceptual loss. L1 loss quantifies the absolute difference between the restored image n' and the ground-truth n. Perceptual loss evaluates the variance between feature maps extracted from the ground-truth image n and the restored image n', utilizing a pre-trained VGG-19 network before activation.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

- Adversarial Loss: Focused on restoring realistic textures.
- Facial Component Loss: Prioritizes the enhancement of facial details.
- Identity Preserving Loss: Consideration of maintaining identity during the restoration process.
- These learning objectives collectively guide the GFP-GAN training to produce realistic, detailed, and identity-preserving facial restorations.

$$L_{rec} = \lambda_{l1} \|n' - n\|_1 + \lambda_{per} \|\phi(n') - \phi(n)\| \quad (6)$$

The values of λ_{l1} and λ_{per} represent the weights of L1 and perceptual loss, respectively. It is of utmost importance to recognize that these values perform a vital function in determining the overall excellence of the reconstructed image. Accordingly, it is vital to choose them carefully to achieve optimal results.

- Adversarial loss: Adversarial loss is a loss function that is utilized in generative models to motivate the generated output to resemble the real data. In the case of GFP-GAN, this function aims to generate genuine textures and solutions that belong within the natural image manifold. Logistic loss is the specific form of adversarial loss that GFP-GAN utilizes, just like the one utilized in StyleGAN2 [4].

$$L_{adv} = -\lambda_{adv} E_{n'}[\mathrm{softplus}(D(n'))] \quad (7)$$

where λ_{adv} represents the weight of the adversarial loss and D denotes the discriminator. The softplus function is utilized to guarantee that the discriminator output is positive, which is required for the logistic loss function. In this manner, the adversarial loss assists GFP-GAN in generating images n' that are difficult to distinguish from genuine images n by the discriminator. The discriminator and generator were trained consecutively in a contradictory mode to create the most exceptional and feasible output.

- Facial Component Loss: To enhance the facial components that have significant perceptual value, we propose a novel method that involves the loss of facial components with local discriminators for the left eye, right eye, nose and mouth. The proposed technique involves cropping the areas of interest using ROI align. We then train individual and small local discriminators for each region to differentiate between restored patches that are real and those that are not. This approach aids in bringing the patches closer to the natural facial component distributions, resulting in a higher quality restored image n'. The objective of this technique is to enhance the fidelity of the restored facial details while preserving the face's identity. The purposed methodology may have the potential to improve the quality of facial reconstruction, particularly in situations where facial features are missing or damaged. $L_{comp} = \sum_{ROI} [\lambda_{local} E_{n'_{ROI}} [\log(1 - D_{ROI}(n'_{ROI}))]] + \lambda_{fs} \|\mathrm{Gram}(\psi(n'_{ROI})) - \mathrm{Gram}(\psi(n_{ROI}))\|$ (8)

The second technique employed is the feature style loss, which is founded on the acquired discriminators and endeavors to equate the Gram matrix statistics of genuine and restored patches. The Gram matrix computes the feature correlations and efficiently captures texture information. The features are derived from various layers of the acquired local discriminators, and the Gram statistics of intermediate representations from the genuine and restored patches are matched. The feature style loss outperforms the prior feature matching loss when it comes to creating realistic facial details and diminishing unpleasant artifacts. The facial component loss is defined by a discriminative loss and a feature style loss, whereby the loss weights of local discriminative loss, as well as feature style loss, are represented by λ_{local} and λ_{fs}, respectively.

- Identity Preserving Loss: Identity preserving loss is a term that has been derived from previous work and is quite similar to what is known as 'perceptual loss'. This particular loss is defined on the basis of the feature embedding of a given input face. The feature embedding is obtained by means of a pre-trained facial recognition model called ArcFace [36].

$$L_{id} = \lambda_{id}\|\eta(n') - \eta(n)\| \quad (9)$$

Here, η refers to the face feature extractor, which is ArcFace in this particular case. The weight of the identity preserving loss is denoted by λ_{id}. This model happens to capture the most significant features that are required for identity discrimination. The purpose of identity preserving loss is to ensure that the restored result has a very small distance when compared to the ground truth in the compact deep feature space. The sum of the following losses represents the model's ultimate objective:

$$L_{total} = L_{rec} + L_{adv} + L_{comp} + L_{id}. \quad (10)$$

It is crucial to acknowledge that the primary objective of this loss function is to ensure the preservation of identity on the occasion of image restoration.

G. Overview of Real-ESRGAN

Real-ESRGAN [1] stands out as a robust and efficient approach for enhancing low-resolution images through deep learning methods for upscaling. Our proposed pipeline ensures seamless implementation, involving the preparation of input images, processing through the Real-ESRGAN model, and efficient tiling for large images. To enhance user experience, we've integrated utility functions for deep network interpolation and parallelized output. The utilization of Real-ESRGAN has resulted in substantial improvements in the quality and details of low-resolution images, offering users an impressive outcome. This underscores its effectiveness and value as a tool for image enhancement.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

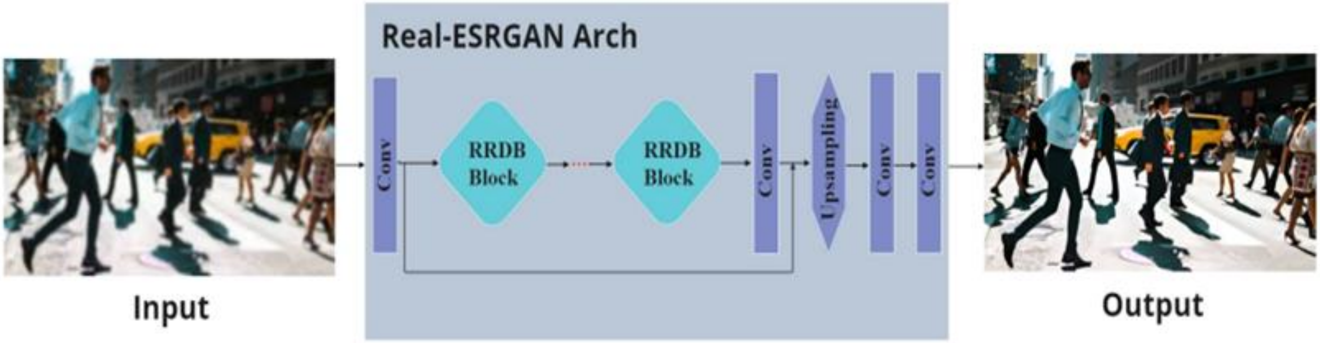

Fig 3: Overview of Real-ESRGAN Framework

H. ESRGAN Generator:

We adopted the generator architecture from ESRGAN [17], employing a deep network featuring multiple residual-in-residual dense blocks (RRDB) [37]. This generator is dedicated to achieving super-resolution with a scale factor of ×2. The process commences with the input image, traversing the RRDB blocks in a forward-feed manner. Each RRDB block comprises multiple convolutional layers with residual connections, enabling the model to capture and enhance fine-grained details. The gradual refinement of the input image through iterative iterations enables the model to perfect high-frequency components, resulting in an image of superior quality and detail. Following the RRDB blocks, the output undergoes additional processing through a series of convolutional layers, responsible for generating the ultimate super-resolved image. These layers aim to enhance the characteristics in the input image and align them with the target resolution. Consequently, the super-improved image achieved through this process exhibits significantly higher resolution than the initial low-resolution input image, with vastly enhanced characteristics. The utilization of convolutional layers in this manner represents a crucial advancement in generating super-resolution images [17].

IV. DATASETS AND IMPLEMENTATION

A. Training Dataset

Our model has been trained on the FFHQ collection [3], which comprises 70,000 images of exceptional quality. As part of the training procedure, all images are transformed to retain a 512x512 pixel resolution. Our model has undergone training on synthetic data that approximate to low-quality images and demonstrates generalization to real-world images during the inference process. We have employed a degradation model, comparable to the one utilized in a previous study [27].

$$m = (n / k\sigma) \downarrow r + n\delta \cdot \text{JPEG}q \tag{11}$$

The notation $k\sigma$ indicates that the standard deviation of the Gaussian blur kernel is a stochastic variable randomly chosen from a range spanning 0.2 to 10. The numerical value of $k\sigma$ determines the level of blurring applied to the image, serving as a parameter that regulates the extent of this blurring. The variable "r" represents the down-sampling factor, causing a reduction in image size by a random factor ranging from 1 to 8. The degree of resolution reduction is directly proportional to the value of "r," with higher values indicating a greater reduction. The degree of δ is randomly selected from a range between 0 to 15, where higher values signify an elevated level of additive white Gaussian noise introduced into the image. δ represents the extent of noise incorporation. The quality factor q, used in JPEG compression, is randomly assigned values between 60 and 100. Higher values of q correspond to superior quality in the JPEG-compressed image.

- Testing Dataset: Our model underwent testing on the Celeb A dataset [38], consisting of 200K images. Notably, this dataset has no identity intersection with Q, ensuring the overall integrity and reliability of our model. The images in this dataset are diverse and complex, showcasing a wide array of degradation from minor to severe. Some images, particularly old photos, exhibit severe degradation in both details and color, posing a significant challenge to our model. For the synthetic test dataset, we randomly selected 1% of images from the Celeb A dataset.

B. Implementation

In this investigation, we adopt the widely-used StyleGAN2 [4] as our generative facial prior, renowned for its effectiveness in generating facial images with a resolution output of 512x512. To maintain a compact model size, we configure the channel multiplier of StyleGAN2 to one. Our model incorporates a U-Net for degradation removal, featuring seven down samples and seven up samples, each equipped with a residual block [39]. To further enhance the model's capabilities, we introduce a modulation network for each CS-SFT layer. This network, comprising convolutional layers that generate affine parameters (α, β), contributes to refining the model. The training mini-batch size is set at 12, and we augment the training data with horizontal flip and color jittering for improved robustness. To ensure precision, we focus on perceptually significant components – the left eye, right eye, nose, and mouth, identified as key facial elements. Utilizing ROI align [40] with face landmarks from the original training dataset, we crop each component. Our implementation of Real-ESRGAN builds upon the successful ESRGAN architecture, incorporating RRDB blocks and additional convolutional layers for super-resolution. The models are implemented using the PyTorch infrastructure and trained on four NVIDIA Tesla P40 GPUs, providing the computational resources needed for optimal results in our study.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

V. RESULTS AND COMPARISONS

A. Experimental Result

The synergistic potential of amalgamating Real-ESRGAN and GFPGAN methodologies comes to fruition, unveiling its prowess in image restoration and enhancement. Evident in the ensuing figures are captivating experimental images, eloquently portraying the extraordinary strides made in elevating the visual appeal and authenticity of the restored images The transformative impact of this collaborative framework is poignantly encapsulated in the visual narratives, where the amalgamated methodologies unfold as virtuoso artists sculpting nuanced details and revitalized clarity onto the canvas of degraded images. The figures stand testament to the dynamic interplay of these methodologies, heralding a new era in the realm of image processing where the boundaries of restoration and enhancement are expansively redefined.

Fig 4: Illustrates the Input and Output Images of Noisy Images

Fig 5: Illustrates the Input and Output Images of Blurry Images

Fig 6: Illustrates the input and output images of scratched images

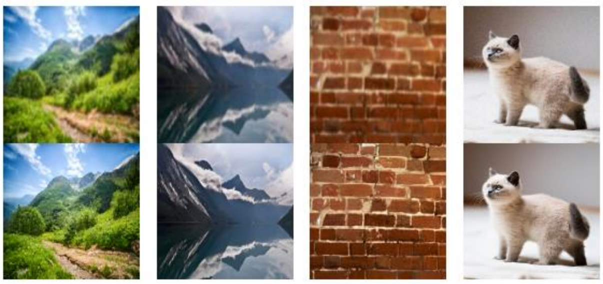

Fig 7: Illustrates the Input and Output Images for Detailed Enhancement and Artifact Removal

- **Evaluation Matrices:** The CelebA-Test with Ground-Truth (GT) was utilized to assess the performance of our model, and three commonly used metrics were employed, namely non-reference perceptual metrics (FID), Perceptual metric (LPIPS), and Pixel wise metrics (SSIM). The *Frechet* Inception Distance (FID) score was 19.5747, indicating a reasonably good match between the generated pictures and the ground-truth pictures. A lower score on the FID indicates higher quality and similarity to actual images. The Learned Perceptual Image Patch Similarity (LPIPS) value was 0.2545, indicating a favorable perceptual similarity between the generated images and real images. A lower value on the LPIPS indicates better perceptual quality. Structural Similarity Index Measure (SSIM) score was 0.9575, reasonable structural similarity between the produced images and ground-truth images. Higher SSIM values signify better structural similarity.

B. Comparison with State-of-the-art Method

In our study, we compared our model to the state-of-the-art face restoration method GFP-GAN. We utilized three metrics that are widely accepted to evaluate the effectiveness.

Table-1: Quantitative Comparison on CelebA-Test

Models	FID	LPIPS	SSIM
GFP-GAN (Original Model)	24.9205	0.2945	0.9546
GFP-GAN with REALESRGAN (Proposed Model)	19.5747	0.2545	0.9575

In our study, we compared our model to the state-of-the-art face restoration method GFP-GAN. We utilized three metrics that are widely accepted to evaluate the effectiveness.

The quantitative outcomes of our model in each condition are presented in Table 1. It is worth mentioning that these metrics were computed based on a synthetic test dataset. For this dataset, we randomly sampled 1% of the images from the larger CelebA dataset. This selection process ensures a representative evaluation of the models' performance on a diverse set of facial images. Our model exhibited a lower LPIPS in this scenario, implying that it outperformed the original model. Additionally, our model achieved a lower FID, indicating that the generated outputs were closer to the actual face distribution and natural image distribution. Notably, our method holds a better identity than the original model when considering pixel-wise metrics such as SSIM, as it has slightly higher degree of identity retention. These findings suggest that our proposed method is superior to the state-of-the-art GFP-GAN in terms of its ability to restore faces while retaining identity and achieving a closer resemblance to real face and natural image distributions. Qualitative results are presented in Fig.4. Thanks to the powerful generative facial prior, GFPGAN is capable of restoring faithful details in various facial features such as the eyes, teeth, nose, and more. Additionally, Real-ESRGAN restores texture details, resulting in more realistic face and texture restorations for real-world samples.

Retrieval Number: 100.1/ijitee.B979213020124
DOI: 10.35940/ijitee.B9792.13020124
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In contrast to other methods, which either struggle to remove degradations effectively or introduce unnatural textures and artificial faces that are not faithful to the original image, our approach achieves a superior balance between removing degradations and preserving the authenticity and realism of the restored images. 2) Our model displays a unique property in that it does not require CUDA extensions, permitting it to function on both GPUs and CPUs. The previous model was only GPU-friendly. 3) The inclusion of specialized convolutional layers such as ModulatedConv2d and StyleConv in our StarGan model brings additional capabilities and flexibility. These layers permit more precise regulation of the transformation and modulation of characteristics, resulting in improved outcomes. Conversely, the original StarGAN model lacks these specialized layers, potentially limiting its ability to capture and manipulate complex image features. 4) To enhance performance and stability during training, our model utilizes a Sequential model with specific layers such as NormStyleCode, Linear, and LeakyReLU activation for the style MLP. The NormStyleCode layer normalizes the style code, the Linear layer performs transformations, and the LeakyReLU activation function prevents the vanishing gradient problem. This design choice leads to improved results. In contrast, the original model employs a regular Linear layer without these additional components, which may limit its capacity for capturing diverse style variations and achieving optimal training dynamics.

C. Limitations

Our model has restricted capabilities in detecting particular facial features, specifically the eyes, nose, and mouth. While our model has the potential to reduce the visibility of scratches in images, it does not guarantee complete eradication. The complicated and varied nature of scratches present a challenging obstacle to complete removal.

VI. CONCLUSION AND FUTURE WORKS

Our collaborative framework is a significant evolution in the field of image restoration. It brings together the strengths of Real-ESRGAN and GFP-GAN methodologies, resulting in unparalleled outcomes in terms of enhancing image resolution, reducing noise, eliminating artifacts, and restoring lost details. Our experiments have shown exceptional visual quality and accuracy when compared with existing techniques, which is noteworthy. Moreover, this framework possesses vast potentialities for practical use in various fields, and it acts as a durable base for future investigations into the realm of image restoration. In conclusion, our technique epitomizes a considerable and innovative contribution to the domain of image restoration, an approach that can potentially revolutionize the way we deal with this crucial area of research. In the future, our objective is to increase the potentialities of our model by including advanced techniques to enhance the eradication of scratches in images. In addition, we intend to improve the model's facial identification and analysis capabilities by increasing its coverage to incorporate more facial components such as eyebrows, ears, and chin. These developments will result in more effective and accurate research result

DECALARION STATEMENT

Funding	No, I did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material/ Data Access Statement	Reference[3]
Authors Contributions	Mousumi Hasan: Developed the initial research idea. Designed methodology and experimental approach. Drafting initial version of the paper and analysis the experimental results. Conducted data analyses . Nusrat Jahan Nishat:: As a second author she Actively participated in the research process, including data collection, analysis, and interpretation. Contributed to the review of relevant literature in the field. Tanjina Rahman: Offered insights into prior research that informs the current work. Assisted and wrote the literature review. Mujiba Shaima:Assisted in writing, data pre-processing, implementing and editing revised paper . Quazi Saad ul Mosaher: Collaborated with the authors to coordinate the manuscript preparation and result analysing and presenting. Mohd. Eftay Khyrul Alam: Created visual representations of data and assisted in analysis project.

REFERENCES

1. X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-esrgan: Training real-world blind super-resolution with pure synthetic data," in International Conference on Computer Vision Workshops (ICCVW).
2. X. Wang, Y. Li, H. Zhang, and Y. Shan, "Towards real-world blind face restoration with generative facial prior," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021. https://doi.org/10.1109/CVPR46437.2021.00905
3. T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4401–4410. https://doi.org/10.1109/CVPR.2019.00453
4. T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 8110–8119. https://doi.org/10.1109/CVPR42600.2020.00813
5. J. Gu, Y. Shen, and B. Zhou, "Image processing using multi-code gan prior," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 3012–3021.
6. X. Pan, X. Zhan, B. Dai, D. Lin, C. C. Loy, and P. Luo, "Exploiting deep generative prior for versatile image restoration and manipulation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 11, pp. 7474–7489, 2021. https://doi.org/10.1109/TPAMI.2021.3115428
7. S. Menon, A. Damian, S. Hu, N. Ravi, and C. Rudin, "Pulse: Self-supervised photo upsampling via latent space exploration of generative models," in Proceedings of the ieee/cvf conference on computer vision and pattern recognition, 2020, pp. 2437–2445. https://doi.org/10.1109/CVPR42600.2020.00251
8. Y. Guo, J. Chen, J. Wang, Q. Chen, J. Cao, Z. Deng, Y. Xu, and M. Tan, "Closedloop matters: Dual regression networks for single image super-resolution," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 5407–5416. https://doi.org/10.1109/CVPR42600.2020.00545

9. D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, "Non-local recurrent network for image restoration," Advances in neural information processing systems, vol. 31, 2018S. https://doi.org/10.1007/978-3-030-01234-2_18
10. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 286–301. https://doi.org/10.1109/CVPR42600.2020.00243
11. J. Liu, W. Zhang, Y. Tang, J. Tang, and G. Wu, "Residual feature aggregation network for image super-resolution," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 2359–2368. https://doi.org/10.1109/TIP.2017.2662206
12. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," IEEE transactions on image processing, vol. 26, no. 7, pp. 3142–3155, 2017.
13. M. El Helou, R. Zhou, and S. Susstrunk, "Stochastic frequency masking to im- ¨ prove super-resolution and denoising networks," in Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVI 16. Springer, 2020, pp. 749–766. https://doi.org/10.1007/978-3-030-58517-4_44
14. O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "Deblurgan: Blind motion deblurring using conditional adversarial networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8183–8192. https://doi.org/10.1109/CVPR.2018.00854
15. J. Guo and H. Chao, "Building dual-domain representations for compression artifacts reduction," in Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer, 2016, pp. 628–644. https://doi.org/10.1007/978-3-319-46448-0_38
16. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, ´ A. Tejani, J. Totz, Z. Wang et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681–4690. https://doi.org/10.1109/CVPR.2017.19
17. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in Proceedings of the European conference on computer vision (ECCV) workshops, 2018, pp. 0–0. https://doi.org/10.1007/978-3-030-11021-5_5
18. J. Chen, J. Chen, H. Chao, and M. Yang, "Image blind denoising with generative adversarial network based noise modeling," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3155–3164. https://doi.org/10.1109/CVPR.2018.00333
19. Q. Cao, L. Lin, Y. Shi, X. Liang, and G. Li, "Attention-aware face hallucination via deep reinforcement learning," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 690–698. https://doi.org/10.1109/CVPR.2017.180
20. X. Yu, B. Fernando, R. Hartley, and F. Porikli, "Super-resolving very lowresolution face images with supplementary attributes," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 908–917.
21. Y. Chen, Y. Tai, X. Liu, C. Shen, and J. Yang, "Fsrnet: End-to-end learning face super-resolution with facial priors," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2492–2501. https://doi.org/10.1109/CVPR.2018.00264
22. D. Kim, M. Kim, G. Kwon, and D.-S. Kim, "Progressive face super-resolution via attention to facial landmark," arXiv preprint arXiv:1908.08239, 2019.
23. C. Chen, X. Li, L. Yang, X. Lin, L. Zhang, and K.-Y. K. Wong, "Progressive semantic-aware style transformation for blind face restoration," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 11 896–11 905. https://doi.org/10.1109/CVPR46437.2021.01172
24. X. Yu, B. Fernando, B. Ghanem, F. Porikli, and R. Hartley, "Face super-resolution guided by facial component heatmaps," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 217–233.
25. X. Li, M. Liu, Y. Ye, W. Zuo, L. Lin, and R. Yang, "Learning warped guidance for blind face restoration," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 272–289.
26. X. Li, W. Li, D. Ren, H. Zhang, M. Wang, and W. Zuo, "Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2706–2715.
27. X. Li, C. Chen, S. Zhou, X. Lin, W. Zuo, and L. Zhang, "Blind face restoration via deep multi-scale component dictionaries," in Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. Springer, 2020, pp. 399–415. https://doi.org/10.1007/978-3-030-58545-7_23
28. R. Abdal, Y. Qin, and P. Wonka, "Image2stylegan: How to embed images into the stylegan latent space?" in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 4432–4441. https://doi.org/10.1109/ICCV.2019.00453
29.] J. Zhu, Y. Shen, D. Zhao, and B. Zhou, "In-domain gan inversion for real image editing," in Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVII 16. Springer, 2020, pp. 592– 608. https://doi.org/10.1007/978-3-030-58520-4_35
30. K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "Ghostnet: More features from cheap operations," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 1580–1589. https://doi.org/10.1109/CVPR42600.2020.00165
31. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.
32. Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, "Dual path networks," Advances in neural information processing systems, vol. 30, 2017.
33. X. Zhao, Y. Zhang, T. Zhang, and X. Zou, "Channel splitting network for single mr image super-resolution," IEEE transactions on image processing, vol. 28, no. 11, pp. 5649–5662, 2019. https://doi.org/10.1109/TIP.2019.2921882
34.] X. Wang, K. Yu, C. Dong, and C. C. Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 606–615. https://doi.org/10.1109/CVPR.2018.00070
35. T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu, "Semantic image synthesis with spatially-adaptive normalization," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 2337–2346. https://doi.org/10.1109/CVPR.2019.00244
36. J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4690–4699. https://doi.org/10.1109/CVPR.2019.00482
37. H. Li, "Image super-resolution algorithm based on rrdb model," IEEE Access, vol. 9, pp. 156 260–156 273, 2021.
38. Z. Liu, P. Luo, X. Wang, and X. Tang, "Large-scale celebfaces attributes (celeba) dataset," Retrieved August, vol. 15, no. 2018, p. 11, 2018.
39. S. Wu, S. Zhong, and Y. Liu, "Deep residual learning for image steganalysis," Multimedia tools and applications, vol. 77, pp. 10 437–10 453, 2018.
40. K. He, "Georgia gkioxari, et al. mask r-cnn," arXiv preprint arXiv:1703.06870, 2017.
41. Akila, Mrs. P. G., Batri, K., Sasi, G., & Ambika, R. (2019). Denoising of MRI Brain Images using Adaptive Clahe Filtering Method. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 1s, pp. 91–95). https://doi.org/10.35940/ijeat.a1018.1091s19
42. Sharma, Dr. K., & Garg, N. (2021). An Extensive Review on Image Segmentation Techniques. In Indian Journal of Image Processing and Recognition (Vol. 1, Issue 2, pp. 1–5). https://doi.org/10.54105/ijipr.b1002.061221
43. Image Synthesis M/2D/HWT in VLSI Technology. (2019). In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 1, pp. 2976–2982). https://doi.org/10.35940/ijitee.a9119.119119

AUTHORS PROFILE

Mousumi Hasan is an Assistant Professor in the Computer Science Department at Bangladesh Army International University of Science & Technology, previously contributing to the software industry and developed projects like human resource management, automatic attendance system and merchandising soft wares.

Holding a B.Sc. and M.Sc. in Computer Science and Engineering from Jahangirnagar University, Bangladesh, she actively pursues research in diverse areas such as image processing, machine learning, data mining, and software engineering. Engaged in projects that integrate OBE curriculum and software engineering, Mousumi has authored several journal papers and conference paper published by IEEE, showcasing her expertise in image processing, machine learning, deep learning, and network security. Her multifaceted contributions underscore her dedication to advancing computer science knowledge.

Quazi Saad ul Mosaher, a Software Engineer at Red Orch Technology Limited, holds a B.Sc. in Computer Science from Bangladesh Army International University of Science & Technology. Passionate about advancing technology, he specializes in image processing, pattern recognition, machine learning, and data mining. With a dedicated commitment to academic discourse, Quazi has contributed significantly to these fields, showcasing his expertise through research publications. His professional journey reflects a keen interest in leveraging innovative solutions to address challenges in computer science, making him a valuable asset in the realm of software engineering and cutting-edge technological advancements. Currently he is involved in projects of student's affairs data mining and network security and steganography.

Mujiba Shaima, an accomplished software developer with an MBA and an ongoing second MSC in computer science at Monroe College, New York, boasts a career spanning since 2011. Her professional journey has been marked by impactful roles in Bangladesh's esteemed banking and pharmaceutical sectors, where she excelled as both a software developer and a quality assurance engineer. Mujiba's commitment to staying at the forefront of technological advancements is evident through her pursuit of further education, showcasing her dedication to expanding her skills and contributing to the ever-evolving field of computer science.she has publised several articles on image processing and data mining with different journals and conferrences.

Tanjina Rahman, an alumna of Bangladesh Army International University of Science and Technology in Cumilla, Bangladesh, is a budding professional who embarked on her journey with a notable internship in the esteemed IT section of a renowned bank. Her academic background and practical exposure showcase her commitment to gaining comprehensive knowledge and hands-on experience in the field of information technology. Tanjina's internship in the bank's IT support section reflects her enthusiasm for applying theoretical insights to real-world challenges, setting the stage for a promising career in the dynamic realm of technology and banking.She is Interested in research in the fileds of machine learning and image processing. Currently she is trying for higher studies aborad.

Nusrat Jahan Nishat, a graduate from Bangladesh Army International University of Science and Technology in Cumilla, Bangladesh, has demonstrated academic excellence and practical experience. Her internship in the IT section of a bank reflects her commitment to gaining real-world insights into the industry. Additionally, Tanjina's noteworthy achievement includes the publication of an article in a conference, showcasing her engagement in scholarly activities and her contribution to the academic discourse. With a solid educational foundation and a proactive approach to professional development, Nusrat Jahan Nishat is poised for success in the dynamic field of information technology.

Mohd. Eftay Khyrul Alam, a graduate of Bangladesh Army International University of Science and Technology in Cumilla, Bangladesh, has showcased practical prowess through his industrial training at the ICT wing of BAIUST. During his training, he successfully developed a Hostel and Dining Management System, which is currently operational at BAIUST's hostel and dining facilities. This accomplishment underscores his ability to apply theoretical knowledge to real-world scenarios, contributing to the efficiency of institutional processes. Mohd. Eftay Khyrul Alam's dedication to enhancing systems reflects his commitment to both academic and practical advancements in the field of technology.