IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 June 2025, accepted 8 July 2025, date of publication 14 July 2025, date of current version 18 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3588956

RESEARCH ARTICLE

Enhancing Image Quality by Optimizing and Fine-Tuning Multi-Fidelity Generative Adversarial Networks

NOOR BAHA ALDIN
Department of Electrical and Electronics Engineering, Hasan Kalyoncu University, 27410 Gaziantep, Türkiye
e-mail: noor.aldin@hku.edu.tr

ABSTRACT Achieving a balance between image quality and computational efficiency is one of the most challenging tasks in image enhancement. Conventional single-fidelity methods focus on structural integrity; yet, they are ineffective in improving perceptual quality, resulting in unrealistic images. The limitations identified in this paper are addressed by introducing a unique Multi-Fidelity Generative Adversarial Network (MF-GAN). The quality of the image is enhanced by the dynamic integration of low-fidelity and high-fidelity models using a composite loss function and dual-generator design. This method achieves high superior perceptual quality and structural integrity while preserving computational efficiency by gradually improving image details. The model was tested on standard natural image benchmarks, including Set5, Set14, and DIV2K, it achieved a Natural Image Quality Evaluator (NIQE) of 4.55, a Peak Signal-to-Noise Ratio (PSNR) of 30.5 dB, and a Structural Similarity Index (SSIM) of 0.92, outperforming existing methods in terms of fidelity and perceptual quality when compared to techniques like ESRGAN (Enhanced Super-Resolution GAN) and RDN (Residual Dense Network). Regarding to the computational efficiency, MF-GAN has 27.6M parameters which is competitive to other models. It achieves an inference time of 1.32 seconds with 203.7G FLOPs (Floating Point Operations). Although MF-GAN is a little faster than RDN, its computational cost is higher than models like ESRGAN and SRGAN (Super-resolution GAN). These results show that MF-GAN is a promising method for image enhancement, since it effectively balances computational efficiency and perceptual image quality.

INDEX TERMS GAN, image enhancement, image restoration, multi fidelity optimization.

I. INTRODUCTION

The main aims of image enhancement are to improve the overall visual quality of images and make them interpretable for both human observers and other systems. Image enhancement finds applications in a wide range of fields due to its essential role in digital image processing [1]. Traditional image enhancement methods often fail to achieve the ideal balance between processing speed and image quality. Resource intensive models could provide exceptional results despite their high processing costs. Low-fidelity models, on the other hand, yield less accurate results but are much faster. An integration of these two methods can achieve better results and at the same time reducing processing costs [2], [3].

This article introduces a novel multi-fidelity optimization method based on Generative Adversarial Networks (GANs). It improves image quality by integrating low-fidelity and high-fidelity models dynamically, balancing accuracy and computational efficiency [4], [5]. This approach enables refinement using high-fidelity models to achieve high results after effective initial improvements using low-fidelity models. The multi-fidelity method generates either perfect or almost ideal solutions more efficiently, as high-fidelity evaluations have reduced processing costs [6], [7]. High-fidelity models are always computationally costly even if they may provide good results. While less exact, low-fidelity models obtain results quickly. Combining the two forms of

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

models allows the multi-fidelity method to close the gap. The study aims to increase the capability of the model to capture improved images. Evaluation results on Set5, Set14, and DIV2K demonstrate that the proposed method performs well on a variety of natural image datasets, which differ in resolution, texture, and structural complexity, but share the same visual domain. Especially in terms of perceptual metrics such as Learned Perceptual Image Patch Similarity (LPIPS) and visual artifact, the results exceed state-of-the-art (SOTA) methods such as RDN and ESRGAN, while maintaining competitive PSNR and SSIM values. Main contributions of the study consist of the following:

- Development in various degrees of the proposed GAN architecture offers realism to enhance image quality and lower processing costs. The low-fidelity model rapidly improves images, while the high-fidelity model refines the outputs to achieve higher quality.
- The model was tested using three benchmark datasets; which are diverse with real world images but are not broad in other domains such as medical or satellite. These datasets show that the suggested method is adaptable and robust in various scenarios. As this study shows, MF-GAN is a good option for applications demanding for high resolution outputs as they balance image quality and computational cost.

II. RELATED WORK

Improvements in optimization methods have resulted in significant theoretical and practical enhancement of image quality. This section contributes to the literature by explaining the specific properties of the multi-fidelity optimization method and evaluating its effectiveness in image analysis and enhancement tasks. The enhanced images methods such as Gamma correction and histogram equalization are methods that have gained popularity over time. Even though these methods effectively increase contrast, their use is often limited in complex scenarios such as varying image quality where manual parameter modifications are required [8].

The development of GANs has led to better images in recent years. One of the earliest techniques to increase resolution was SRGAN, which created a model for low-resolution images [9]. While SRGAN and other GAN methods are often designed to achieve specific objectives, such as color correction or super-resolution, they have limited generalization capability. Despite the uniqueness of GANs, it is important to understand their limitations [9]. Their major focus is on image enhancement therefore they are often limited to specific tasks like noise reduction or color adjustment. Cycle GANs work well for unsupervised image-to-image interpretation, but they struggle to maintain consistency among images with different levels of accuracy [10], [11]. To maintain sharpness and detail, ESRGAN, a refined version of SRGAN, enhanced perceptual loss and residual-in-residual blocks [10]. Despite these improvements, ESRGAN still challenges with images of different quality, which might result in too sharp artefacts. Additionally, despite their effectiveness, RDN models aren't suitable for real-time processing applications because of their higher computational costs [11]. The primary purpose of these early models was super-resolution, modern techniques like DPSR aim to improve tasks such as dual-pixel images, quality, and depth perception accuracy [12]. Similar to earlier methods, it is still designed for certain tasks and has difficulties extrapolating to images with a variety of properties. The conditional GANs provide more controllable in terms of image generation, but it cannot generalize the varying image quality [10], [13].

This study overcomes the limitations of existing methods by providing a framework capable of restoring images at various fidelities. The system's superiority over the results on Sets5, Set14, and DIV2K shows its ability to handle variety of image resolutions and detail levels in the domain of natural image. MF-GAN extends the application range of GANs by including multi-fidelity conditioning instead of conventional GANs, which typically address uniform fidelity. This leads to exceptional performance on the mentioned benchmark datasets [12], and allows it to well adapt to diverse image properties. Theoretically, the proposed architecture can be generalized to other domains, however, this paper focuses on natural image benchmarks. By combining different levels of realism and detail, it effectively uses processing power and provides users with more control over image enhancement. It improves upon previous methods and offers a more complete and useful solution. The model's ability to handle a range of natural image inputs with different characteristics makes it a strong choice for many image improvement tasks outside of the specialty domains that single-fidelity model techniques like ESRGAN, RDN, and DPSR commonly serve.

III. METHODOLOGY

MF-GAN uses both low-fidelity and high-fidelity enhancement models to enhance the quality of the low-resolution images iteratively, which outputs that nearly close to high-resolution images in terms of content and detail. Unlike conventional GANs methods that use a single generator for direct upscaling. This structured approach optimizes image enhancement in two stages:

1. **Low-Fidelity Enhancement** which initially enhances basic features and overall image structure from low-resolution inputs with minimal computational overhead, focuses on basic image reconstruction, reducing pixel-level errors between enhanced low resolution images and their high resolution. The loss function used is Mean Squared Error (MSE), which is considered the best for initial improvements.

$$L_{low} = \frac{1}{WH} \sum_{i=1}^{W} \sum_{j=1}^{H} (I_{ij}^{HI} - G_{low}(I^{LI})_{ij})^2 \quad (1)$$

where I^{HI} is the high-resolution target image, I^{LI} is the low-resolution input image, G_{low} is the low-fidelity model, and W, H are the width and height of the image, respectively.

2. **High-Fidelity Enhancement** refines the outputs from low-fidelity using Residual Dense Blocks (RDBs) to

recover fine textures, it also incorporates more complex loss functions, which include perceptual loss, edge loss, and adversarial components to ensure the final image quality is close to true high-resolution images. By focusing on sharp edges, textures, and structural consistency, it enhances details.

This study suggests a unique GAN model to overcome the difficulties in high-fidelity images without paired datasets. This model uses low and high-fidelity generating systems to enhance image resolution gradually, surpassing typical GAN methods. In the proposed approach, the output of the low-fidelity generator is fed as an intermediary input into the high-fidelity generator, refining the details and textures incrementally. A unique composite loss function that integrates perceptual, adversarial, and total variation components aims to create realistic, structural and detailed images. The final generator's discriminator ensures the accuracy of the image improvements. This prevents over-smoothing while preserving the original quality. The model was evaluated on DIV2K after being trained for 1,000 epochs on Set5 and Set14, demonstrating its scalability and efficiency across the mentioned datasets.

The discriminator that has been trained to determine the accuracy of high-resolution images has a major impact on the refining of the generator output. This discriminator is an essential part of the adversarial training loop as it uses binary cross-entropy loss to push the generators to produce further improvements. The model relies on the exact interaction of the generators and the discriminator to achieve an appropriate balance between preventing over-smoothing, which is typical in GAN-generated images, and maintaining fidelity to the original data. This systematic method guarantees improvements to images, with the low-fidelity model handling simple reconstruction and the high-fidelity model focusing on achieving perceptual similarity to high-resolution images. To demonstrate the effectiveness of the Multi Fidelity (MF) method, each model is essential to the overall process of enhancement. The preprocessed dataset included unpaired images of both high and low resolutions. These images served as inputs for the model. e model performed reliably and consistently throughout a thousand training epochs. Its performance and quality were demonstrated by comparisons between high-resolution and enhanced images. The enhanced image quality was assessed using several metrics, including PSNR and SSIM. Although the results demonstrate strong enhancement capabilities on natural datasets, extensive evaluation will be conducted with domain specific data as future work. The general workflow of MF-GAN is visualized in Figure 1.

A. GENERATOR ARCHITECTURE

The generator architecture is the key component of the MF image enhancement system. The main goal is the improvement of low-resolution images by different processing. This system follows a dual-model strategy, integrating low and high-fidelity techniques to balance between computational

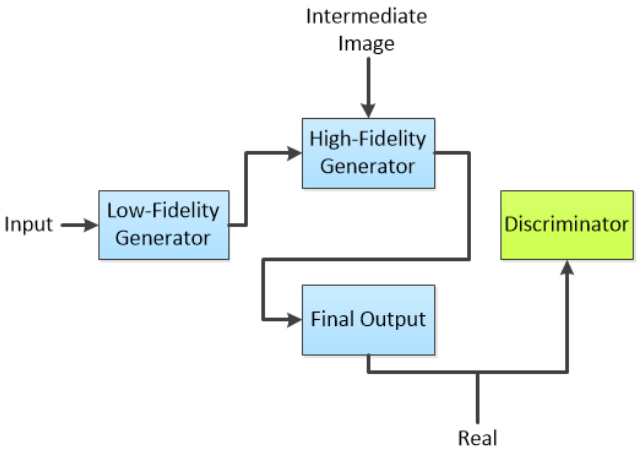

FIGURE 1. General architecture of MF-GAN.

efficiency and the quality of perception. This process allows such results to be more content and detail consistent with real High Resolution (HR) images. The low-fidelity model focuses on a fundamental structural enhancement during upscaling, at a low computational cost. It has 32 convolutional layers with 3×3 kernels to capture significant features from the input image. The 3×3 kernels in the low-fidelity model are designed to balance computational efficiency and feature extraction capability. Smaller kernels allow the model to capture fine details and yet maintain the number of parameters. The output's original spatial dimensions are preserved by last convolution layer with three filters of same size. Conv layers apply a weighted sum and an activation function, like Rectified Linear Unit (ReLU) and Sigmoid to enhance the image. The ReLU adds nonlinearity which makes the model able to realize its patterns, whereas Sigmoid helps smooth out pixel values to look more natural. The high-fidelity model uses the low fidelity model's intermediate image to improve texture, restore finer details, and increases perceived quality. Each feature input x is multiplied by the appropriate weight to improve the quality of the image. During the training process, these weights are modified to enable the model to focus on the most relevant features, including textures and edges, to produce a clearer and more detailed output. The inputs' weights are calculated as follows:

$$y = \sum_{i=1}^{n} (\omega_i . x_i) + b \tag{2}$$

The bias term b is an additional parameter that allows the model to shift the activation function, which helps the generator in learning more accurate feature representations. The activation function determines the output of the neuron. Three key activation functions are used, the first is ReLU:

$$f(x) = \begin{cases} x, & if\ x \geq 0 \\ 0, & if\ x < 0 \end{cases} \tag{3}$$

This function helps in preserving important image details by keeping positive activations, and avoid the problem of

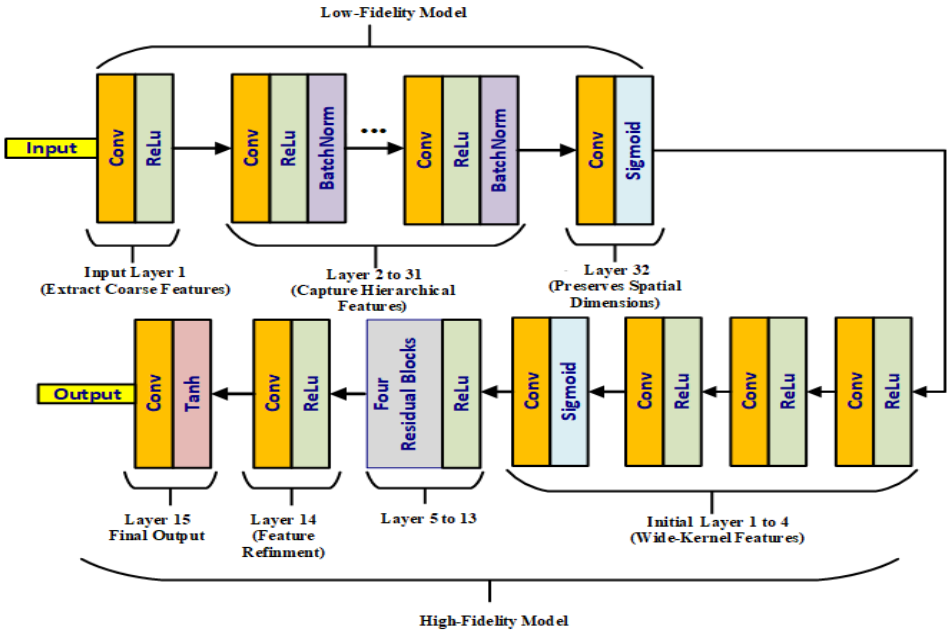

FIGURE 2. Layered architecture of the MF-GAN generator.

vanishing gradient. The second nonlinear function is the Sigmoid activation:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

This function is used to enhance image's pixel values. After the image passes through the weighted summation and activation functions, the output is an intermediate enhanced image. This serves as input to the high-fidelity model, where fine details and textures are further improved. Finally, the Tanh activation function is used in the output layer of the high-fidelity model to normalize pixel values, making the output more realistic:

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5}$$

To enhance clarity, the full layered structure of the generator used in MF-GAN is shown in Figure 2. The generator follows a two-path model architecture which consists of a low-fidelity path designed for the coarse enhancement and a high-fidelity path designed for perceptual and textural refinement.

The design of the generator is outlined in Table 1, 2.

TABLE 1. Low fidelity generator architecture.

Layer	Type	Kernel Size	Filters	Activation	Description
1	Conv	3×3	32	ReLU	Input Layer
2–31	Conv	3×3	32	ReLU, BatchNorm	29 Layers
32	Conv	3×3	32	Sigmoid	Output layer

The first convolutional layer of the high-fidelity model consists of 64 filters of 9×9 kernels, used to extract broad features. Residual blocks play a key role in preserving fine details in the image. Residual blocks are important for preserving fine details structure on the image. The high-fidelity model, contains residual blocks and is important for maintaining fine details and avoiding information loss for further feature extraction. Skip connections within these blocks support the direct flow of gradients, addressing the vanishing gradient problem and enhancing the training stability. Convolutional layers, each with 64 filters and 3×3 kernels, are included in each block. These layers help extract important features while maintaining the integrity of the original image. Before reaching the final layer, the image is passed through additional convolutional layers with 32 filters (3×3 kernels) and a final layer with 32 filters (5×5 kernels) that enhance the image even more. Skip connections allow information to bypass some layers, preventing the loss of fine details, and the output passes through the ReLU activation function inside residual blocks to maintain depth and nonlinearity. The Tanh activation in the last layer scales the pixels to the range of -1 and 1, making the result photo-realistic.

The model uses a combined loss function during training to optimize image quality, including perceptual loss, which uses the features from a pre-trained Visual Geometry Group (VGG19) model for deep content understanding [13], the edge loss can help to enhance the edge property to make the image more clear [14], adversarial loss used to guarantee that the output image can deceive the discriminator to identify it as a high-quality real image [15], [16], and total variation loss encourages smoothness and reduces occurrence of pixilation [17]. The generator's architecture

TABLE 2. High fidelity generator architecture.

Layer	Type	Kernel Size	Filters	Activation	Description
1-4	Conv	9×9	64	ReLU, Sigmoid	Initial wide kernel
5–13	Residual Blocks	3×3	64	ReLU	4 blocks
14	Conv	3×3	32	ReLU	Feature refinement
15	Conv	5×5	32	Tanh	Final output layer

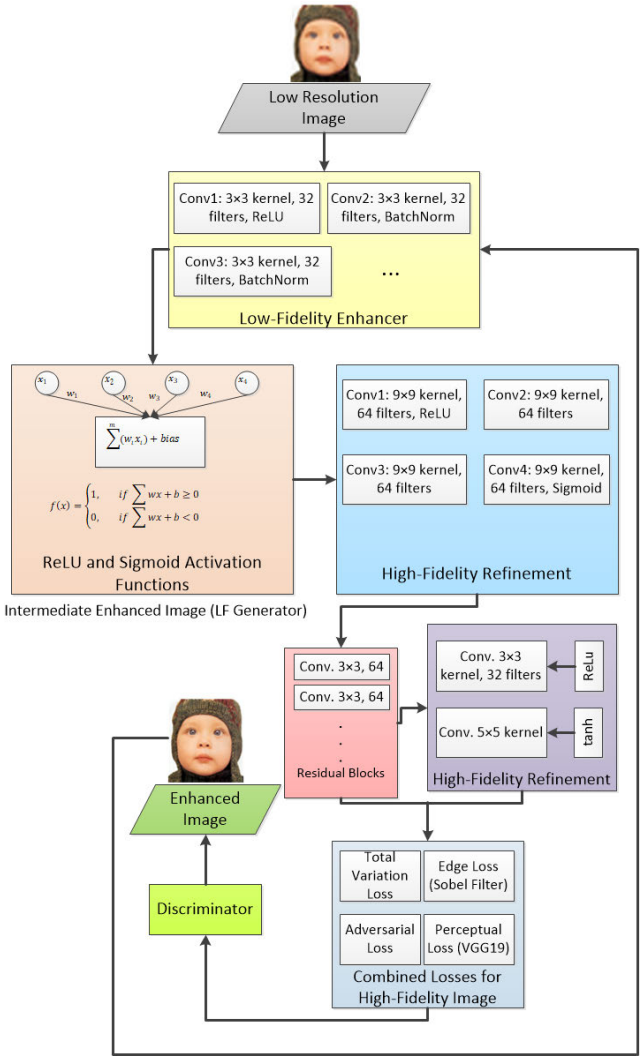

FIGURE 3. Dual-model architecture of MF-GAN.

thoroughly enhances low-resolution images, improving both structural fidelity and perceptual realism. Figure 2 shows the dual-model architecture of the MF image enhancement, illustrating the sequential stages from the low-fidelity to the high-fidelity model and their interaction with the discriminator.

Conv Layer: Convolutional Layer, **ReLU Activation**: Non-Linearity **BatchNorm**: Normalization, **Edge Loss:** Sobel Filter Enhancing Edges

B. DISCRIMINATOR ARCHITECTURE

The MF image enhancement method in the GAN architecture depends critically on the discriminator. It ensures perceptual realism at all levels by learning from hierarchical outputs of the two generative models via adversarial feedback. Its objective is to test how photo-realistic these high-fidelity models are and confirm if they are indistinguishable from real high-resolution images. The discriminator is designed as a Convolutional Neural Networks (CNNs) which is particularly built to assess the authenticity of the enhanced images. Its architecture consists of multiple convolutional layers that use 32 filters of size 3 × 3 kernels. After that, it continues on to a second layer that uses 64 filters of sizes 5 × 5 and 9 × 9 kernels. The discriminator uses kernels of multiple sizes (3 × 3, 5 × 5, and 9 × 9) to capture hierarchical features, from details refinement to broader patterns Strides of 2 and ''same'' padding are used by both layers to reduce spatial dimensions while keeping major textural and structural, at the same time, enabling efficient feature extraction. LeakyReLU was chosen as the activation function to avoid the dying ReLU problem and keep the flow of gradient flow during the training. A dropout of 0.4 rate helps avoids overfitting by randomly deactivating neurons, and helps the network to generalize better [18], [19].

This method works by disabling a few neurons temporarily, thereby increasing the robustness of the model [20]. The newly obtained output from the convolutional layer is flattened and is fed through the dense layer to create a single output unit with a sigmoid activation function. This unit helps the discriminator's binary classification task, which predicts the likelihood whether the processed image is real or generated by the high-fidelity model. The discriminator provides adversarial feedback to the generator by predicting whether an image is real. This conditioning sets up a positive feedback loop that pushes the generator to create more realistic images, thereby balancing the trade-off between fidelity and perceptual quality. The architecture of the discriminator is shown in Table 3. The adversarial feedback loop between generator and discriminator allows to improve the output and create realistic image enhancements of high quality at the same time [19].

Figure 4 shows the sequentially extracts of the input data from low-to-high level features, applies dropout for regularization, and terminates with a dense sigmoid layer for binary classification.

C. NOVELTY OF MF-GAN IN IMAGE ENHANCEMENT

The MF-GAN is proposed with a dual-generator architecture which combines low-fidelity and high-fidelity models in a sequential enhancement process. MF-GAN provides a significantly different approach by incrementally improving image details gradually, by going through various levels to

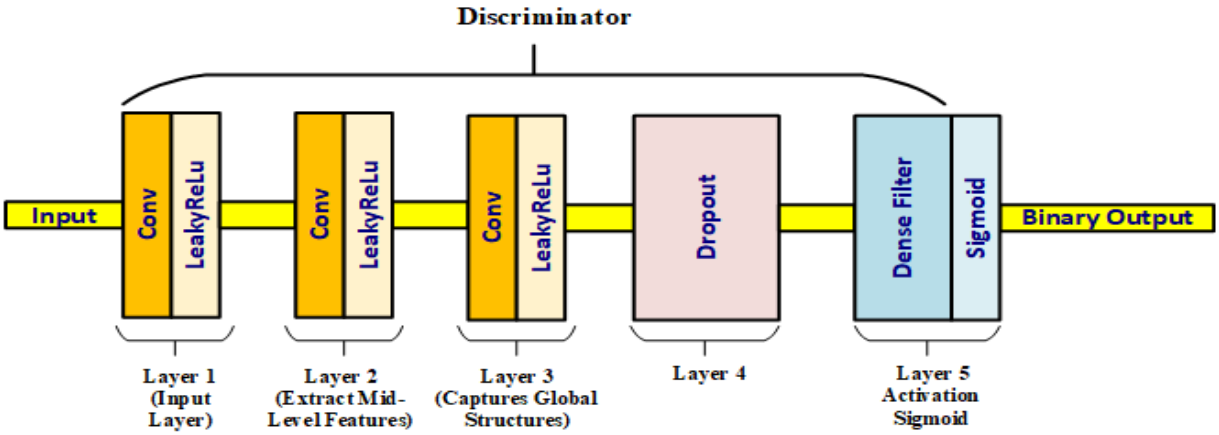

FIGURE 4. Layered discriminator architecture of MF-GAN.

TABLE 3. Discriminator architecture.

Layer	Type	Kernel Size	Filters	Activation	Description
1	Conv	3×3	32	LeakyReLU	Input layer
2	Conv	5×5	64	LeakyReLU	
3	Conv	9×9	64	LeakyReLU	
4	Dropout	–	–	–	0.4
5	Dense	–	1	Sigmoid	Binary output

find balance between efficiency and the quality of image. This is quite different from traditional GAN designs, which rely on a single generator-discriminator structure to upscale images in one step. On the other hand, current techniques such as ESRGAN, RDN and DPSR have shown good results in improving images, they mostly focus on one-stage upscaling, which results in the loss of fine-grained features or excessive sharpening artifacts.

A key innovation of MF-GAN lies also in its hybrid loss function with adaptive weighting, which combines perceptual loss to maintain content similarity, adversarial loss to make images more realistic, edge preservation loss to keep sharpness and avoid blurring, and total variation loss to reduce noise and ensure smoother transitions in the enhanced images. These components combined to enhance sharpness, structural integrity, and natural image texture. Adaptive weight configurations ($\alpha = 1.0$, $\beta = 1$, $\gamma = 0.5$, $\delta = 0.5$, $\epsilon = 1.0$) are used to optimize the perception distortion, and they are fixed throughout training. This ensures improved perceptual realism and structural similarity.

MF-GAN exhibits strong scalability and generalization, achieving good results across natural image datasets, such as Set5, Set14, and DIV2K. For instance, the MF-GAN even outperformed the ESRGAN (29.1dB, 180.1G FLOPs) and RDN (29.0 dB, 210.3G FLOPs) in the DIV2K dataset with the PSNR of 30.5 dB along with 203.7 G FLOPs. In comparison to RDN and DPSR, MF-GAN increases inference time slightly 1.54% compared to RDN (1.30s) but reduces inference time by 1.49% compared to DPSR (1.34s). It requires 13.1% more computations than ESRGAN but achieves higher output quality, while requiring 3.14% fewer computations than RDN and 3.28% fewer than DPSR (210.6G FLOPs). This method provides an effective balance between computational efficiency and the fidelity of the output by using low-fidelity approximations as a basis, which are then improved with high-fidelity refinements to enhance the finer details. The proposed method not only maintain the original image quality, but also ensures scalability and adaptability in natural image datasets.

D. LOSS FUNCTION

The generator and discriminator in the MF image enhancement system are trained using different loss functions. The evaluation for the enhanced images of several aspects, including edge preservation, overall image quality, and likeness to original images, depends on these functions.

E. PERCEPTUAL LOSS

The perceptual loss method is used to compare the enhanced images with the ground-true high-resolution images in terms of content and differences. It makes use of a previously trained VGG19 network to extract feature representations at various levels from both generated and actual images [21]. The loss is calculated using the Euclidean distance between the feature maps of the actual and generated images.

$$L_{perc} = \sum_{l \in L} \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\emptyset_l \left(I^{HR} \right)_i - \emptyset_l (G_{high}(I^{LR}))_i \right)^2 \tag{6}$$

where G_{high} is the high-fidelity model, $\emptyset_l$ denotes the feature map extracted by the VGG19 model at layer l, L is the set of selected layers from VGG19 used for feature extraction, and N_l is the total number of elements in the feature map at layer l.

The weight of perceptual loss improves perceptual quality by matching high level features, increasing it in case if the generated image lacks realistic textures. Decreasing it if there are artifacts in the generated image [22].

F. EDGE LOSS

The edge loss contributes to preserve edge characteristics, which is one of the most important for visual perception in high-resolution images, and prioritizes the sharpness and clarity property of edges in the generated images with respect to the original images [23]:

$$L_{edge} = \frac{1}{M}\sum_{i=1}^{M}\left|Sobel(I^{HR})_i - Sobel(G_{high}\left(I^{LR}\right))_i\right| \tag{7}$$

The Sobel () function applies the Sobel operator to an image to detect edges by highlighting regions of high spatial frequency. M is the number of all the pixels in the image.

G. ADVERSARIAL LOSS

Adversarial loss, which is essential for the GAN formulation, ensures the generated images to seem real to the discriminator. It measures how well the discriminator is fooled by the generated images [24]:

$$L_{adv} = -\log(D\left(G_{high}(I^{LR})\right)) \tag{8}$$

where $D\left(G_{high}(I^{LR})\right)$ is the discriminator output to generate image indicating the probability of the image is real.

The weight is adjustable to help the generator produce more realistic images; if the GR image seems distorted or unclear, it can be increased; if it becomes unstable or noisy, it can be decreased [25].

H. TOTAL VARIATION LOSS

Total variation loss is added to maintain the spatial smoothness of the generated images, reducing noise and helping achieve a more coherent and natural appearance [26].

$$L_{tv} = \sum_{i,j}\left(\left|Y_{i,j} - Y_{i+1,j}\right| + \left|Y_{i,j} - Y_{i,j+1}\right|\right) \tag{9}$$

where $Y_{i,j}$ represents the pixel values at position (i, j) in the generated image. The weight of it reduces noise and improves spatial smoothness, increasing it if the GR image is noisy, decreasing it if the GR image loses fine details.

I. COMBINED LOSS FOR HIGH-FIDELITY MODEL

The final loss function combines all of the previously discussed loss functions and is weighted based on how much each one contributes to the production of realistic, high-quality image improvements.

$$L_{combined} = \alpha L_{prec} + \beta L_{edge} + \gamma L_{adv} + \delta L_{tv} \tag{10}$$

where α, β, γ, δ are the weights assigned to each loss component, determining their impact on the overall loss. These loss functions are critical for training the MF image enhancement models, ensuring every aspect of image quality like sharpness and realism is finely adjusted. By balancing these loss functions, the model learns to produce images that look real, maintain visual clarity, and retain precise details. Figure 3 shows the step-by-step training of the GAN framework's generator and discriminator, which begin with low-resolution inputs and refine them into high-resolution outputs.

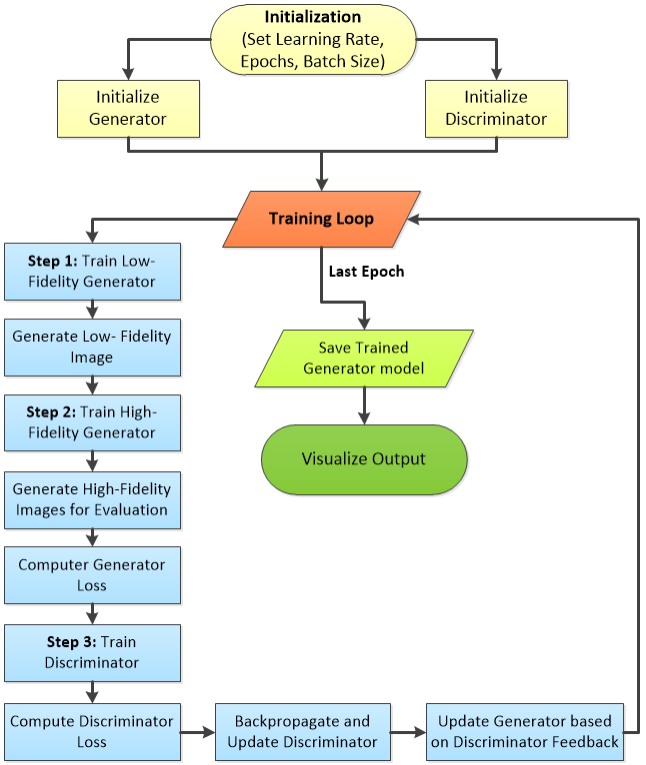

FIGURE 5. Training process for multi-fidelity GAN.

IV. EXPERIMENTATION

A. DATASET

For the validation, the study is tested on more complex and high-resolution images from DIV2K, in addition to the Set5 and Set14 datasets used for training and testing, offering a more challenging benchmark to evaluate the model's performance. High quality images first down sampled to low-resolution for the validation of the enhancement methods. Its small size makes testing and development easier, allowing for rapid evaluation of improvements. Additionally, the variety of content within images also makes it feasible to evaluate other aspects of the image, such as sharpness and clarity.

B. EXPERIMENT SETTINGS

This research utilized datasets, which included several images downsampled from high resolution to low resolution [27]. Several image enhancement techniques were applied to test these images. The low-fidelity model used in the experiment contained convolutional layers for feature enhancement and

early image upsampling. This specific model was developed especially for the test. The high-fidelity model improved these developments and generated outputs with high-density images by using many residual blocks. The discriminator model has to be tested for quality to differentiate augmented from real images [28]. In this study, a batch size of 32 images was used to train both models. To increase the size of the training set, various data augmentation techniques were used. The learning rate was set at 1e-4 and then adjusted via a decay schedule to improve convergence and minimize overfitting [29]. Both low and high-fidelity discriminator and generator models provide updates during training to help stabilize the adversarial training. This method maintains balance between the generator and discriminator, reducing the possibility of model collapse or the discriminator dominating the training dynamics [30].

C. EVALUATION METRICS

The effectiveness of the image enhancement models was evaluated using the following metrics in image processing; PSNR is a widely used quantitative metric to measure the quality of enhanced images. It compares the similarity between a high-resolution original image and the enhanced image:

$$PSNR = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \tag{11}$$

where MAX_I is the highest pixel value of the image, and MSE indicates the mean squared error value of the enhanced and original images [31].

The perceived quality of the enhanced images is evaluated using SSIM. It takes into account changes in structural information, luminance, and contrast, while PSNR focuses on absolute errors. The SSIM metric is mathematically defined as:

$$SSIM\ (x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{12}$$

where μ and σ are the local means, variances, and covariance of the images x and y, respectively. The constants c_1 and c_2 are introduced in order to stabilize the process in the presence of weak denominators to achieve numerical stability [32].

NIQE is a no-reference index which estimates the perceptual image quality by using natural scene statistics [33]. The following arithmetic expression is used to calculate the value of NIQE:

$$NIQE = \sqrt{(\mu_t - \mu_p)^T \sum_p^{-1} (\mu_t - \mu_p)} + trace(\sum_t + \sum_p - 2(\sum_t \sum_p)^{\frac{1}{2}}) \tag{13}$$

where μ_t and $\sum_t$ *denote* the mean vector and covariance matrix of the test image features, and μ_p and $\sum_p$ *are* the mean and covariance matrix of the pristine image features [33], [34].

The average squared difference between the original and improved images is measured by MSE. It is calculated as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(I_i^{HR} - I_i^{LR})^2 \tag{14}$$

where N refers to the number of total pixels in the image, I_i^{HR} denotes the high-resolution reference image, and I_i^{LR} refers to the low-resolution input image [35].

The average absolute differences between the original and the enhanced images are determined using the Mean Absolute Error (MAE). It is calculated using the formula:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|I_i^{HR} - I_i^{LR}\right| \tag{15}$$

The square root of MSE of the loss function is Root Mean Squared Error (RMSE). This measure captures the average magnitude of the errors between predicted and observed values, given by:

$$RMSE = \sqrt{MSE} \tag{16}$$

LPIPS, a cutting-edge metric, measures visual similarity by analyzing deep features extracted from a pre-trained network.

The distance of LPIPS between the enhanced image and high-resolution image is calculated as:

$$LPIPS\left(I^{HR}, G\left(I^{LR}\right)\right) = \sum_{l\in L}\frac{1}{N_l}\sum_{i=1}^{N_l} w_l||\Phi^l(I^{HR})_i - \Phi^l(G(I^{LR}))_i||_2^2 \tag{17}$$

The results on the DIV2K dataset show that the MF-GAN's can improve high-resolution images, with an LPIPS 0f 0.17; a lower value indicates better visual similarity.

where I^{HR} represents the high-resolution image, $G\left(I^{LR}\right)$ is the enhanced image, $|\cdot|$ denotes the Euclidean distance, L is the set of layers in the pretrained network, and w_l denotes the learned weight for each layer in the network.

By considering luminance, contrast, and structure into account, the Universal Image Quality Index (UIQI) evaluates how comparable the original and improved images are:

$$UIQI = \frac{4\sigma_{xy}\mu_x\mu_y}{(\sigma_x^2 + \sigma_y^2)(\mu_x^2 + \mu_y^2)} \tag{18}$$

where $\mu_x\mu_y$, σ_x, and σ_y are the standard deviations and means of images x and y, and σ_{xy} represents the covariance of original images x and enhanced y [36].

When comparing the enhanced image to the original one, Visual Information Fidelity (VIF) evaluates how well the visual information is preserved [37]. The mathematical expression for the formula is defined as follows:

$$VIF = \frac{\sum_{j=1}^{N}\sum_{k=1}^{K_j}\log_2\left(1 + \frac{g_{j,k}^2\sigma_{x,j,k}^2}{\sigma_{v,j,k}^2}\right)}{\sum_{j=1}^{N}\sum_{k=1}^{K_j}\log_2\left(1 + \frac{\sigma_{x,j,k}^2}{\sigma_{n,j,k}^2}\right)} \tag{19}$$

In the wavelet transform of image, N denotes the number of decomposition levels. K_j denotes the number of spatial locations in the j-th subband. The gain factor, denoted as $g_{j,k}$ simulate how the models the response of the human visual system. The term $\sigma^2_{x,j,k}$, represents the variance of the coefficients in the j-th subband and k-th spatial location for the reference image. $\sigma^2_{v,j,k}$ denotes the variance of the distortion noise in the j-th subband and k-th spatial location $\sigma^2_{n,j,k}$ represents the variance of the visual noise in the j-th subband and k-th spatial location.

D. COMPUTATIONAL COMPLEXITY ANALYSIS

The computational cost of MF-GAN is evaluated using FLOPs, that scale linearly with the number of arithmetic operations per inference time, which is important for real-time applications, and model parameters, which indicate the model's size and memory requirements. Unlike other models, MF-GAN processes images in two successive steps using a dual-generator technique achieve both efficiency and high-quality results:

1. Low-Fidelity Model: minimizes computational overhead by performing lightweight enhancements.

2. High-Fidelity Model: refines details from the low-fidelity output and enhances perceptual quality using advanced loss functions and RDBs.

Computational efficiency is evaluated by key factors, which are FLOPs, and inference time. These metrics measure the operations required per image and are compared across multiple models in Table 4 [38].

TABLE 4. Computational complexity comparison of image enhancement models.

Model	Parameters (M)	FLOPs (G)	Inference Time (sec)
MF-GAN	27.6	203.7	1.32
ESRGAN	16.7	180.1	0.90
RDN	22.3	210.3	1.30
DPSR	22.4	210.6	1.34
SRGAN	1.5	130.2	0.95

The number of FLOPs required for inference is one of the factors that measure computational efficiency. MF-GAN reduces the FLOPs by 3.3% compared with DPSR, while its dual-generator design leads to higher FLOPs than ESRGAN because it splits the computation between the low-fidelity (coarse) and the high-fidelity (refinement) stages.

MF-GAN has 27.6M parameters, which is more than ESRGAN but comparable to RDN and DPSR. However, this is justified by its dual-generator design, which integrates both low-fidelity and high-fidelity enhancement stages. MF-GAN's design ensures efficient processing by splitting computation between coarse and fine-grained enhancements and balancing refinement capacity, despite the higher parameter count.

It achieves an inference time of 1.32 seconds, which is faster than DPSR (1.34 seconds) but slower than RDN (1.30 seconds). For offline or near real-time applications prioritizing perceptual quality over speed, it achieves an efficient per-image processing time of 1.32 seconds, showcasing its adaptability to resource-constrained environments. The model also reduces inference time by 1.54% compared to RDN and 1.49% compared to DPSR, showcasing its adaptability to resource-constrained environments.

The main advantage of MF-GAN lies in its ability to balance computational efficiency with high-quality image enhancement. Rather than simply applying high-fidelity enhancements to low resolution inputs directly, the low-fidelity generator reduces the computational load on the high-fidelity generator by improving coarse details first. This two-stage refinement ensures that the final result is not only computationally efficient, but also obtains high SSIM and perceptual quality.

To test MF-GAN's scalability, it was trained on the DIV2K dataset, which contains high-resolution images. The model is highly scalable, as its FLOPs, and inference time compared to the current methods. Its two generators model make it feasible to be efficiently applied to a variety of natural image enhancement tasks, with balance of computational complexity and the image perceptual quality.

E. HYPERPARAMETER CONFIGURATION AND TRAINING SETUP

To maintain reproducibility the full set of hyperparameter configuration that was used in the training process is given. Adam optimizer was used to train the model with 1e-4 initial learning rate, decaying by 0.1times every 250 epochs for a total of 1000 epochs. 32 batch size was used. The composite loss weights ($\alpha = 1.0$, $\beta = 1.0$, $\gamma = 0.5$, $\delta = 0.5$, $\epsilon = 1.0$) were tuned based on PSNR, SSIM, and NIQE, as shown in Table 5 and Figures 8–12. Data augmentations such as random flips and rotations were applied.

The overall training procedure is outlined in Algorithm1, detailing the two-stage optimization process for the dual-generator architecture.

lr_images: Low-resolution input images, **hr_images**: High-resolution ground truth images, **MSE_loss**: Mean Squared Error loss (pixel-level fidelity), **VGG_loss**: Perceptual loss derived from VGG-19 feature representations, **Sobel_loss**: Edge loss computed using the Sobel operator, **TotalVariation_loss**: Total variation loss for smoothness and noise suppression, **BCE_loss**: Binary Cross-Entropy loss for discriminator training, α, β, γ, δ: Weighting factors for perceptual, edge, adversarial losses, **adv_loss**: Adversarial loss used to promote realism, **generator_opt/discriminator_opt**: Optimizers for updating the generator and discriminator models, **eval_interval**: Evaluation interval to measure performance, **Dropout**: Regularization technique applied to the discriminator to prevent overfitting.

The full hyperparameter configuration in both generator and discriminator of the MF-GAN architecture is listed in Table 5 for training reproducibility and clarity.

Algorithm 1 MF-GAN Training Procedure

Input: lr_images, hr_images, Generator and Discriminator weights, Hyperparameters α, β, γ, δ
Output: Trained low and high fidelity generators and discriminator
Initialize low-fid_generator, high-fid_generator, and discriminator
Initialize optimizers generator_opt, discriminator_opt
for epoch = 1 to epochs **do**
  for each batch in dataloader **do**
    (lr_images, hr_images) ← preprocess(batch)
    low_output ← low_fid_generator(lr_images)
    low_mse_loss ← MSE_loss(low_output, hr_images)
    Update low_fid_generator with low_mse_loss
    high_output ← high_fid_generator(low_output)
    adv_loss ← adversarial_loss(discriminator(highoutput), real_labels)
    perceptual_loss ← VGG_loss(high_output, hr_images)
    edge_loss ← Sobel_loss(high_output, hr_images)
    tv_loss ← TotalVariation_loss(high_output)
    total_loss ← α·perceptual_loss+β·edge_loss+γ·adv_loss +δ·tv_loss
    Update high_fid_generator with total_loss
    real_pred ← discriminator(hr_images)
    fake_pred ← discriminator(high_output)
    d_loss_real ← BCE_loss(real_pred, real_labels)
    d_loss_generated ← BCE_loss(generated_pred, generated_labels)
    discriminator_loss ← d_loss_real + d_loss_generated
    Apply dropout to discriminator
    Update discriminator with discriminator_loss
  end for
  if epoch mod 200 = 0 **then**
    adjust_lr(generator_opt, factor = 0.9)
    adjust_lr(discriminator_opt, factor = 0.9)
  if epoch mod eval_interval = 0 **then**
    evaluate_metrics(high_output, hr_images)
  end if
end for

F. RATIONALE FOR LOSS COMPONENTS AND WEIGHTING STRATEGY

The weights of the composite loss function were achieved by performing a grid search on the datasets. It was aimed to balance pixel-level accuracy, perceptual realism, structural fidelity, and edge sharpness. The chose the following components based on their complementary objectives; where MSE Loss (α) ensures reconstruction fidelity and stability during training, which is especially important for the early stages of the low-fidelity generator. Perceptual Loss (β) aligns deep feature representations based on VGG-19 and improves the high-level content realism. Edge Loss (ε) is aimed to sharpen boundaries directly, without losing fine details on textures and contours. Adversarial Loss (δ) which guides the generator to generate realistic images indistinguishable, promoting naturalness. Total Variation Loss (γ) reduces noise and pixelation, enhancing the smoothness and continuity. The particular weight settings ($\alpha = 1.0$, $\beta = 1.0$, $\gamma = 0.5$, $\delta = 0.5$, $\varepsilon = 1.0$) were chosen via empirical tuning. Ablation and sensitivity analysis also confirmed that these values yielded the optimal compromise among the image quality metrics; such as PSNR, SSIM, and NIQE. The last configuration was tested on Set5, Set14, and DIV2K, demonstrating consistent improvements over alternative solutions as shown in Table 6.

TABLE 5. Generator and discriminator hyperparameters for MF-GAN training.

Hyperparameter	Generator	Discriminator	Tested Range
Optimizer	Adam	Adam	[1e-5, 1e-3]
Learning Rate	1e-4	1e-4	[1e-5, 1e-3]
Epochs	1000	1000	[500, 1500]
Batch size	32	32	[16, 64]
Kernel Sizes	3x3 (LF), 9x9,3x3, 5x5 (HF)	3x3, 5x5, 9x9	Used for various features extraction
Filters	32 (LF), 64 (HF), 32 (Output)	32, 64	[16-128]
Activation Functions	ReLU, Sigmoid, BatchNorm, Tanh	LeakyReLU, Sigmoid	ReLU, LeakyReLU, BatchNorm, Tanh
Loss Functions	MSE, Perceptual (VGG), Edge, Adversarial	Binary Cross Entropy (BCE)	-
Loss Weights	α=1.0, β=1.0, γ=0.5, δ=0.5, ϵ=1.0	-	[0.1-2.0] for weighted terms
Dropout	-	0.4	[0.2-0.6]
LeakyReLU Slope	-	0.2	[0.1-0.3]
Data Augmentation	Randon flips, rotations		
Training/Validation	80% training, 20% validation	80% training, 20% validation	Consistent split for both models
Stride	-	2	[1,2]
Regularization	L2	L2	Applied to avoid overfitting

TABLE 6. Optimal weight configuration.

Weight	Tested range	Optimal value	Effect on deviation
α	0.5-1.5	1.0	α <1.0: Blurry textures; α>1.0: oversmoothing
β	0.5-1.5	1.0	β<1.0: Blurred edges; β >1.0: Edge artifacts
γ	0.2-1.0	0.5	γ<0.5: Oversmoothing; γ>0.5: Adversarial artifacts
δ	0.2-1.0	0.5	δ<0.5: Noise; δ>0.5: Oversmoothing
ε	0.1-1.0	1.0	ε<1.0: Reduced realism; ε>1.0: Training instability

V. RESULTS

The performance of the MF image enhancement model was evaluated and compared with several existing methods using PSNR, SSIM, and LPIPS metrics. The results are summarized in the Tables 7, 8, and 9 for the Set5, Set14, and DIV2K datasets, and Figures 3–5, respectively.

MF-GAN provides the most visually realistic results, with natural textures, sharp edges and at the same tine minimal artifacts. Both realism and detail recovery are crucial factors

TABLE 7. Quantitative comparison of image enhancement methods based on PSNR, SSIM, and LPIPS metrics (Set5).

Method	PSNR(dB)	SSIM	LPIPS
MF-GAN	31.4	0.92	0.12
ESRGAN	29.6	0.89	0.14
RDN	30.8	0.89	0.17
Conditional GAN	29.7	0.89	0.16
DPSR	29.7	0.90	0.18
SRGAN	29.3	0.90	0.16
EDSR	30.5	0.89	0.19
CycleGAN	28.1	0.88	0.21
Gamma Correction	27.0	0.86	0.24
Histogram Equalization	26.5	0.85	0.24

TABLE 8. Quantitative comparison of image enhancement methods based on PSNR, SSIM, and LPIPS metrics (set14).

Method	PSNR(dB)	SSIM	LPIPS
MF-GAN	30.8	0.90	0.15
ESRGAN	29.3	0.89	0.17
RDN	30.4	0.87	0.20
Conditional GAN	29.0	0.87	0.19
DPSR	28.8	0.88	0.20
SRGAN	29.1	0.89	0.18
EDSR	30.2	0.87	0.19
CycleGAN	28.3	0.80	0.27
Gamma Correction	27.6	0.83	0.26
Histogram Equalization	27.0	0.82	0.34

TABLE 9. Quantitative comparison of image enhancement methods based on PSNR, SSIM, and LPIPS metrics (div2k).

Method	PSNR(dB)	SSIM	LPIPS
MF-GAN	30.5	0.88	0.17
ESRGAN	29.1	0.87	0.19
RDN	30.1	0.86	0.20
Conditional GAN	28.9	0.85	0.21
DPSR	28.8	0.86	0.22
SRGAN	28.6	0.86	0.19
EDSR	30.0	0.85	0.23
CycleGAN	27.6	0.83	0.25
Gamma Correction	26.5	0.82	0.28
Histogram Equalization	25.7	0.81	0.32

in image enhancement. Although ESRGAN generates visually appealing images with realistic textures, complex regions sometimes appear with over sharpened edges. RDN produces clean but oversmoothed images, struggling to recover fine textures, optimal for situations where pixel-level accuracy takes precedence over visual realism. SRGAN generates unrealistic textures in complex scenes, despite its balance between detail and speed. It is suitable for real-time applications in which minor artifacts are acceptable. EDSR is capable of producing smooth results, however, it performs poorly in recovering small details. This is most likely due to the fact that it has limitations in terms of training and dataset mismatches. DPSR can handle arbitrary degradation, but it is less appropriate for high quality enhancement since it yields blurry outputs with limited detail recovery. While the Gamma Correction and Histogram Equalization not only often distort colors and amplify noise, but also failing to restore important features.

FIGURE 6. Visual comparison of image enhancement methods on Set5.

The results of this study show that MF-GAN performs better across all mentioned datasets. It produces the lowest LPIPS values, and the highest PSNR and SSIM. This result indicates that the model has been able to enhance these lost perceptual aspects and structural details compared to the original high-resolution images, as shown in Table 7. By achieving a PSNR of 31.4 dB, SSIM of 0.92, and LPIPS of 0.12 on Set5. This is supported by visual comparisons in Figure 4, which show that MF-GAN retains sharp features and intricate textures in zoomed regions such as the eye/forehead and cheek. While ESRGAN and RDN provide high metric values, they often introduce minor artifacts and over smoothing, particularly in regions with intricate textures. Traditional techniques like Histogram Equalization and Gamma Correction performed the worst, which failed to preserve details and resulted in outputs that were either overly smoothed or blurred.

MF-GAN's performance is shown by its results on the Set14 dataset in Table 8, which achieved a PSNR of 30.8 dB, an SSIM of 0.90, and an LPIPS of 0.15. Unlike ESRGAN, RDN, and other methods that failed to preserve fine details, MF-GAN was able to maintain sharp edges and detailed textures. This is especially obvious in the zoomed areas in Figure 5, where MF-GAN consistently gives clearer and more detailed outputs. Similarly, the DIV2K dataset was used to assess the performance of MF-GAN and other methods; Table 9 provides a summary of the quantitative results. MF-GAN has the best PSNR (30.5 dB), the lowest LPIPS (0.17), and highest SSIM (0.88), indicating best structural integrity and perceptual quality. Visual comparison of a common sample image from the DIV2K database is shown in Figure 6, which demonstrates the differences in texture accuracy, per-

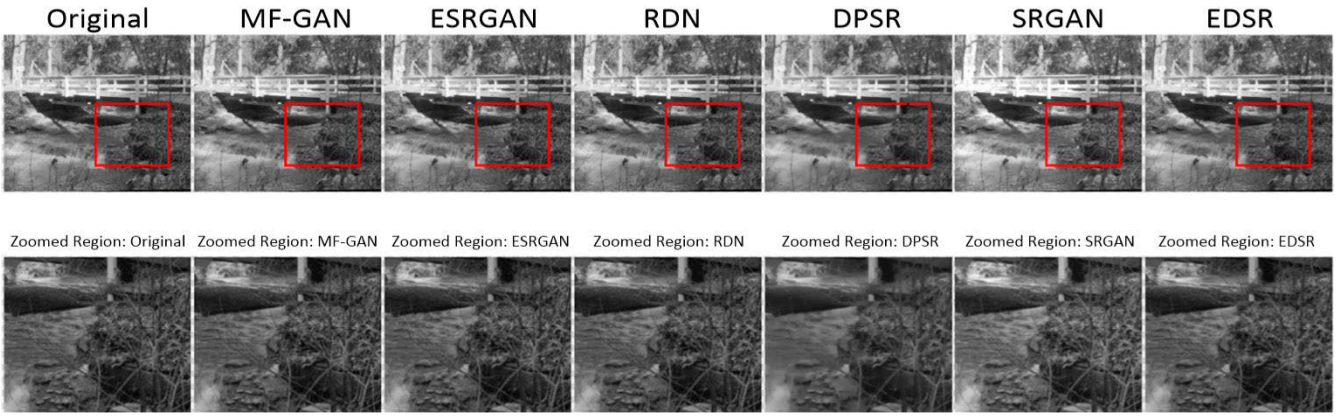

FIGURE 7. Visual comparison of image enhancement methods on sample image from Set14 dataset.

ceptual quality and detail preservation. These images were chosen to represent common challenges in image enhancement, such as complex structures and fine textures. Compared to ESRGAN, which generates less oversharping images, MF-GAN achieves better textures preservation. The fur structure remains natural, avoiding excessive smoothing. RDN's outputs appear oversmoothed; the image is clear but slightly blurry, indicating oversmoothing, it has a lower SSIM (0.86) and higher LPIPS (0.20). Conventional methods, such as Gamma Correction and Histogram Equalization, usually fail to retrieve the important features because of their lower PSNR (26.5 dB and 25.7 dB, respectively) and higher LPIPS (0.28 and 0.32). Figure 6 shows how better MF-GAN is in quantitative and qualitative evaluations. The classical methods often failed, consistently exhibiting lower PSNR and SSIM values and greater LPIPS values. These results highlight their inability to retain image details and structural integrity. MF-GAN minimizes perceptual artefacts while achieving superior texture and structural detail restoration by dynamically combining low and high fidelity stages of processing. MF-GAN's consistent performance across standard benchmark datasets demonstrates its scalability and robustness within natural image enhancement tasks, which therefore makes it a potential solution for applications that demand high perceptual.

The NIQE values for several image enhancement methods, averaged across the Set5, Set14, and DIV2K datasets, are shown in Figure 7. The effectiveness of models was evaluated using NIQE in terms of perceived quality and naturalness.

Lower NIQE values are indicative of better perceptual quality; enhanced images with values less than 5 are considered to reflect good performance. MF-GAN achieves the lowest NIQE value of 4.55, showing its superior performance. This enhanced performance is due to its ability to use both models inside a single framework. The low-fidelity generator handles the initial upsampling. Meanwhile, the high-fidelity generator refines intricate details, ensuring both structural precision and visual realism that align closely with high-resolution outputs. This method produces images that are good not only in quantitative metrics but also in overall visual appeal and naturalness. ESRGAN's slightly higher NIQE (4.61) reflects its oversharpened edges. The higher value (4.72) of RDN indicates that it has a propensity to oversmooth images, reducing fine textures and perceived naturalness. DPSR (4.83) and EDSR (4.91) have the highest NIQE values, reflecting their pixel-level accuracy. As shown in Figure 6, MF-GAN achieves the lowest NIQE value, highlighting its perceptual quality advantage over other methods.

TABLE 10. GAN model quantitative analysis.

Method	PSNR(dB)	SSIM	NIQE
Low-fidelity GAN	25.10	0.820	5.60
High-fidelity GAN	26.85	0.870	5.20
MF-GAN (Full Model)	30.05	0.92	4.55

A. PERFORMANCE EVALUATION OF MF-GAN

The performance of MF-GAN models was evaluated using PSNR, SSIM, and NIQE metrics. The research aims to determine each model's ability to improve low-resolution images, with a particular focus on the advantages of combining high-fidelity and low-fidelity models.

The low-fidelity GAN significantly enhances the basic quality of low-resolution images, as illustrated in Table 10. Its smaller structure, however, limits its ability to consistently capture minute information, thereby producing lower PSNR and SSNR values but a higher NIQE. Despite the model's quality enhancements, it challenges to restore fine details and soft textures such as hair or foliage. High-fidelity GANs aim to improve performance by capturing finer details. The level of image naturalness and texture quality is greatly improved by this level of complexity, which turns into the enhancement of the visual quality. Visual quality is enhanced by the high level of texture quality and image naturalness made possible by this complexity. The results demonstrate the model's best performance on typical samples, showing its potential under the best situation.

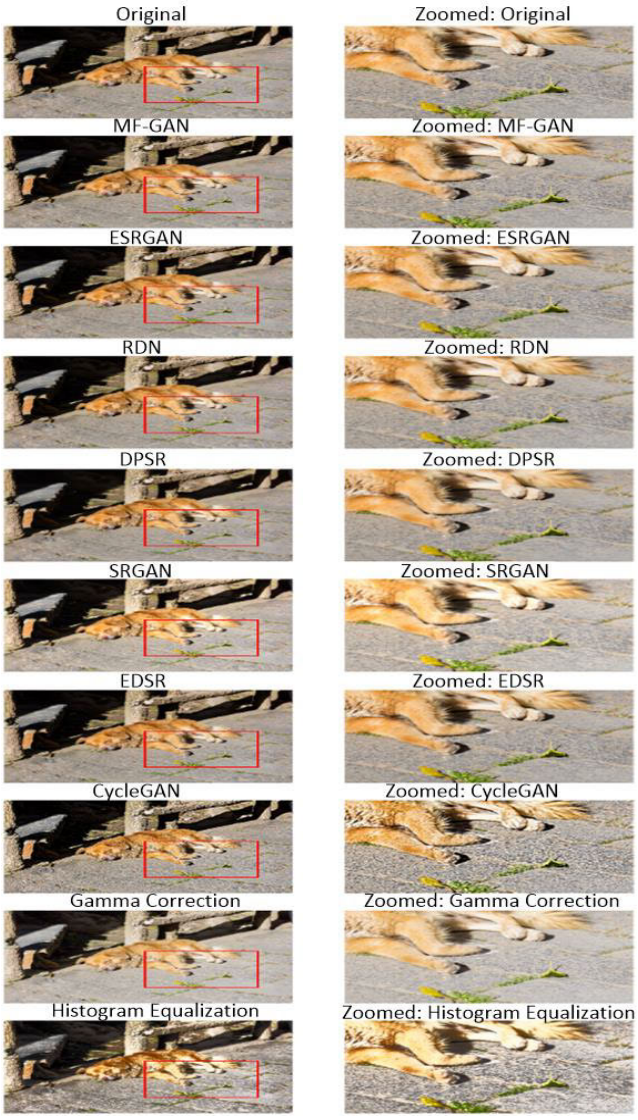

FIGURE 8. Visual comparison of image enhancement methods on DIV2K dataset.

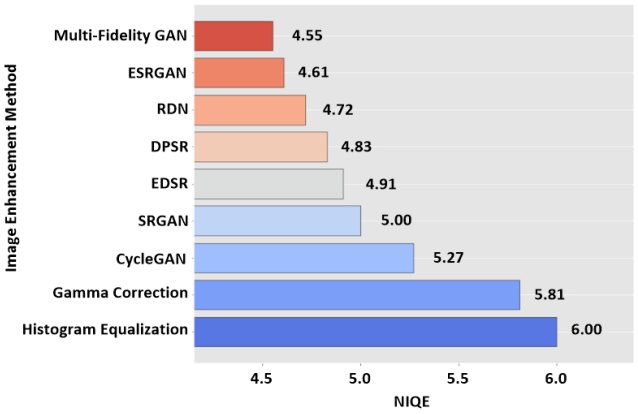

FIGURE 9. NIQE comparison across methods.

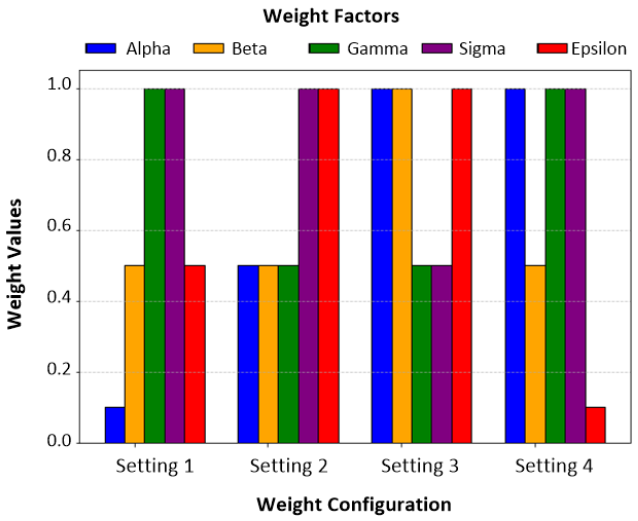

FIGURE 10. Weight values for different settings.

B. IMPACT OF WEIGHT CONFIGURATION ON IMAGE QUALITY

The quality of the enhanced images produced by MF-GAN is largely depends on the weight factors provided by different components of the loss function. Adjusting these weights results in significant changes in PSNR, SSIM, and NIQE values, the corresponding weight values for each setting are shown in Figure 8. Setting 3 outperforms the other tested settings, as it yields the optimal PSNR value and leads to the best overall image quality. The SSIM is at its maximum value when the enhanced image corresponds closely to the structural details of the reference image. It is also observed from the no-reference measure that the lowest value of NIQE is the best in terms of perceptual quality. The best PSNR and SSIM are obtained with the enhanced configuration setting 3, which verifies that this setting keeps structural details well while maintaining high perceptual quality. Setting 2 also offers a stable trade-off, resulting in moderate SSIM and PSNR values, making it an appropriate choice for general cases. Although setting1 maintains a reasonable level of performance, it shows lower values, indicating that it may not be the best choice for all scenarios. Setting 4, however, leads to a decline in both PSNR and SSIM, while NIQE increases, indicating that overweighting adversarial components leads to undesirable perceptual artifacts. The weight configuration is selected to achieve a balance between perceptual realism (NIQE), structural similarity (SSIM), and pixel-level accuracy (PSNR). It affects image quality rather significantly. Regarding maintaining both quantitative and qualitative image fidelity, setting3 provides the best general performance [39], [40].

To balance between pixel-level accuracy and perceptual quality, the composite loss function of the MF-GAN model consists of a few components, each with a certain factor. The pixel-wise loss is controlled by the weight factor α, that sets pixel fidelity. By decreasing α, while still keeping other

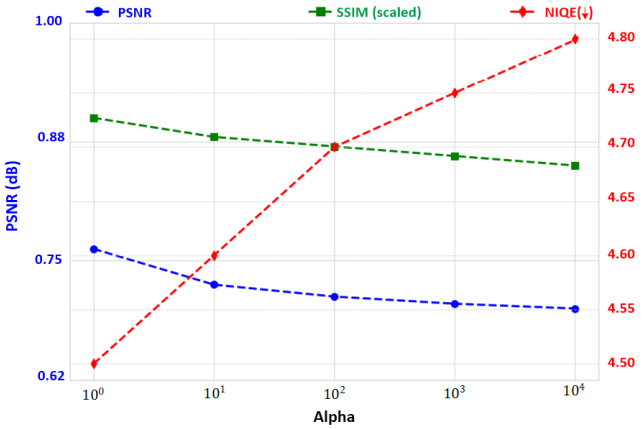

FIGURE 11. Balancing loss weight α for optimal image enhancement.

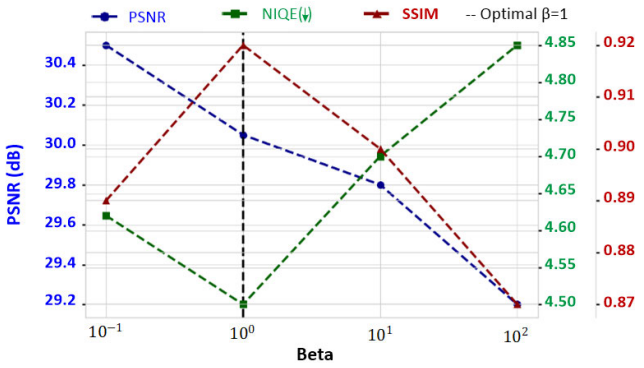

FIGURE 12. Effect of βon image enhancement performance.

components, such as $\beta, \gamma, \delta, \epsilon$, it can give more weights and enhance the role of weights, creating better structural similarity and texture realism while keeping the pixel accuracy acceptable. However, overly low α values lead to overemphasized textures, which lowers the quality of overall image. When $\alpha = 1.0$, the optimal configuration is achieved, which avoids over-smoothing but maintains sufficient pixel-level supervision to avoid artifacts. This value strikes the best balance between pixel-level accuracy and perceptual quality, achieving peak performance across key metrics. Increasing α over the ideal value significantly decreases performance; PSNR reduces by 2.5 dB, SSIM decreases by 0.05, and NIQE decreases by 0.30. Slight increases in α lead to large quality reduction and the benefit tends to be diminishing as the value of the parameter becomes larger. Whereas, the weight factors $(\beta, \gamma, \delta, \epsilon)$ are essential to optimize the loss function. In addition to α, the impact of varying α on image quality metrics is illustrated in Figure 9. When all loss components are balanced, the optimal setting is obtained by $(\alpha = 1.0, \beta = 1, \gamma = 0.5, \delta = 0.5, \epsilon = 1.0)$. This demonstrates the necessity of a balanced approach in the loss function design, ensuring high-quality outputs that excel in both fidelity and perceptual realism output.

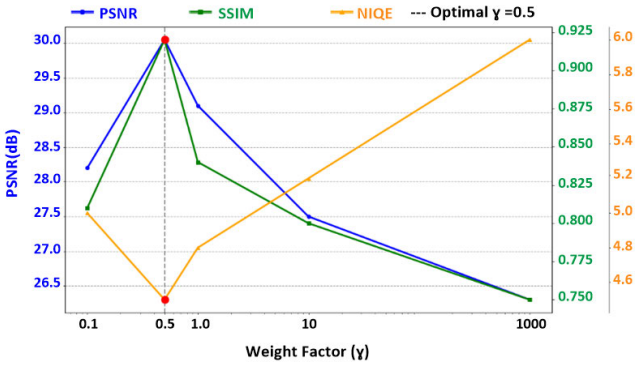

FIGURE 13. Effect of weight factors γ on image quality metrics.

Figure 10 depicts how PSNR, SSIM, and NIQE vary with the loss weight parameter β (perceptual loss) in image enhancement. Results demonstrate the importance of fine tuning of β to enhance the quality of images ensuring the best possible balance between perceptual quality and structural accuracy. The model achieves the best balance between all metrics when β is adjusted to 1.0. Specifically, it achieves the lowest NIQE value of 4.50, suggesting that the enhanced images appear more natural and perceptually realistic; an SSIM of 0.92, demonstrating superior structural similarity; and a PSNR of 30.05 dB, indicating excellent pixel-level accuracy. However, as β grows larger, both PSNR and SSIM decrease, indicating a worse pixel-wise fidelity and structural consistency. Meanwhile, with the higher NIQE values, it means that there is a distortion in perceptual quality, and the images seem less realistic. On the other hand, decreasing β values results in a lower PSNR, indicating an overemphasis on pixel-level accuracy. This indicates the importance of selecting suitable weight tuning for scaling perceptual loss, as well as for preserving structures in an image context. Compared to the previous GAN-based enhancement methods with perceptual and adversarial feedback loops, MF-GAN further includes a dual-stage generator and enhanced loss to better balance image realism and structural fidelity [41], [42].

The impact of tuning the weight factor γ on various key image quality metrics is presented in Figure 11. The x-axis represents the values of γ from 0.1 to 1000, and the y-axis represents how these metrics change with the changing of γ. The best performance is achieved when $\gamma = 0.5$, it avoids over-sharpening and over blurring while ensuring sharp edges, structure consistency, and natural look. The model's inability to preserve edges when γ is less than 0.5 results in blurry edges and overly smooth images that lack realism. On the other hand, when γ is too high $(\gamma > 0.5)$, the model focuses too much on preserving edges, which leads to visual artifacts, unrealistic textures that reduce overall quality, and makes the images look overly sharp. The results demonstrate that $\gamma = 0.5$ is the optimal value, with PSNR and SSIM declining and NIQE rising as γ moves away from this value. These emphasize the importance of tuning

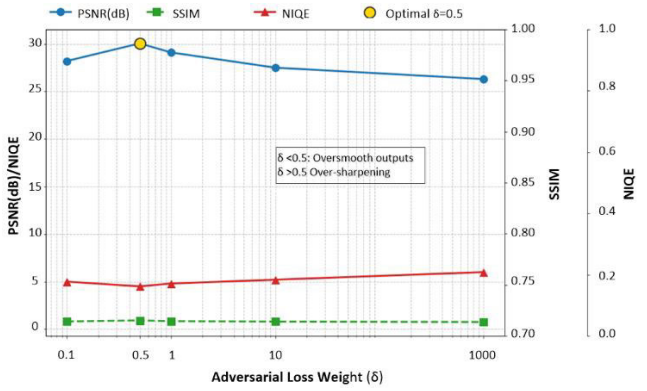

FIGURE 14. Impact of adversarial loss weight (δ) on image quality metrics.

γ to balance PSNR, SSIM, and NIQE, producing visually appealing, coherent, and realistic images.

Figure 12 illustrates the relationship between PSNR, SSIM, and NIQE, demonstrating how three key image quality metrics vary as we adjust the weight factor δ of adversarial loss. The best performance is achieved according to $\delta = 0.5$. This balance results in realistic model with good textures, structural coherence, and realistic appearance while avoiding over-smoothing. The δ is used to control the adversarial loss contribution in the composite loss function, to enforce the generator to produce realistic textures and ensure pixel-level accuracy. When δ is too low ($\delta < 0.5$), then the adversarial loss plays the major role, resulting in overly smooth outputs with poor perceptual quality (higher NIQE). Conversely, when δ is too high ($\delta > 0.5$), the adversarial loss dominates, generating artifacts such as over-sharpening or unnatural textures, resulting in a decrease in all metrics, PSNR drops by 1.4 dB, and NIQE degrades by 0.20. When δ is selected as 0.5, the model achieves an optimal balance that leads the discriminator to effectively drive the texture synthesis process, while preserving both accuracy and realism. This outcome is an image full of clear edges, fine details, and realistic textures. The optimal point is shown in the figure, and it is clear that the performance decreases sharply when δ becomes large. These results show the importance of tuning δ. This result ensures that better images are also structurally consistent with the original content.

Figure 13 examines the impact of changing the weight factors ε (edge loss) on image quality metrics, while keeping other weights fixed. The best balance among the metrics is achieved at $\varepsilon = 1.0$. Edge constraints are underemphasized for the case that $\varepsilon < 1$, resulting in blurred edges and higher NIQE. Conversely, over-prioritizing edges when $\varepsilon > 1$ causes over-sharpening and unnatural textures, decreasing the PSNR by 1.9 dB and degrading the NIQE by 0.10. This balance can be influenced by extreme high or low values, while the moderate ε (1.0) enhances edge sharpness and preserving structures.

Balancing between the loss factors is important for achieving the best image quality in the MF-GAN framework, as depicted in Figure 14. MF-GAN outperforms other

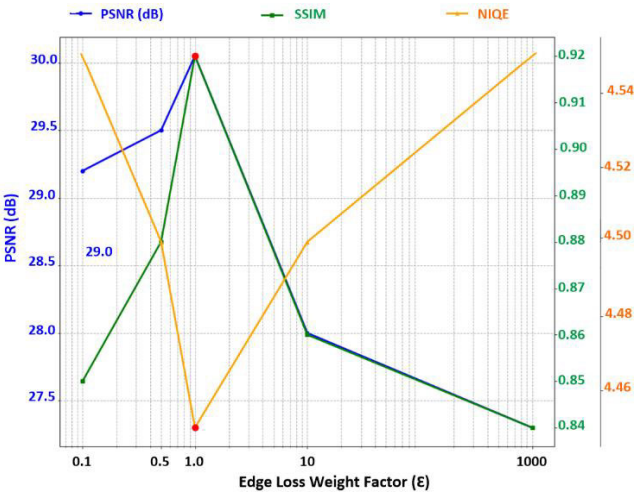

FIGURE 15. Impact of edge loss weight (ϵ) on image quality metrics.

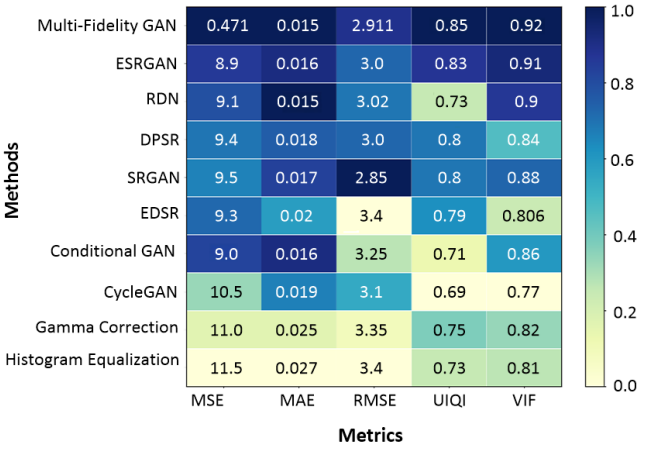

FIGURE 16. Comparison of image quality metrics for image enhancement models.

methods with the lowest MSE (8.471), MAE (0.015), RMSE (2.911), and highest UIQI (0.85) and VIF (0.92), indicating superior accuracy, structural similarity, and visual fidelity. ESRGAN and RDN perform well (ESRGAN: MSE 8.9, VIF 0.91; RDN: MSE 9.1, VIF 0.9) but create noticeable distortions. The results in terms of metrics values and visual quality show that the traditional enhancement methods like Gamma Correction and Histogram Equalization are not suitable, with high MSE values (11.0 and 11.5) and low UIQI scores (0.75 and 0.73), performing poorly, because they do not maintain details and introduce artifacts. Other models like conditional GAN, CycleGAN, and SRGAN perform acceptable results, with SRGAN standing out for its ability to balance generative quality and realism (MSE 9.5, UIQI 0.8). These results illustrate the desired benefit of low and high-fidelity models, which show that MF-GAN's superiority in producing high-quality, structurally coherent image enhancements [34], [35].

In order to demonstrate the convergence of the proposed MF-GAN model, Figure 17 shows the generator and

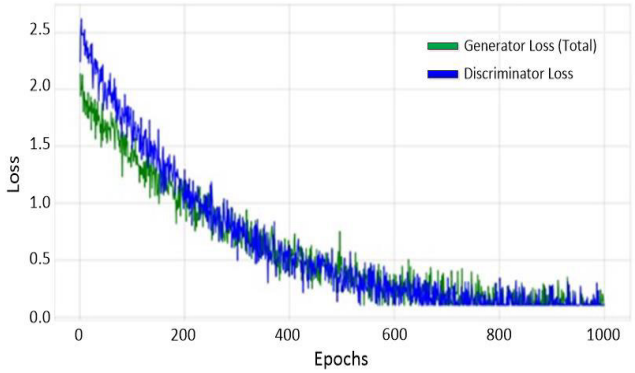

FIGURE 17. Loss convergence during training of MF-GAN.

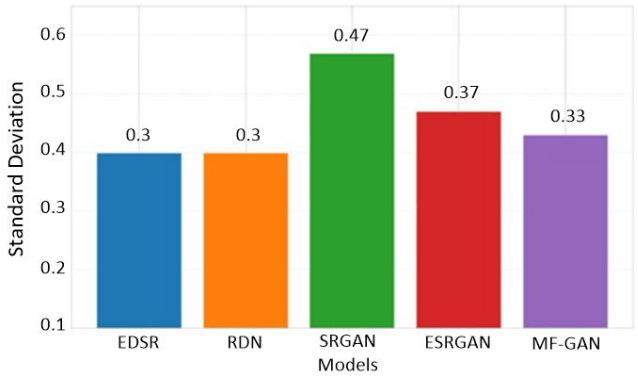

FIGURE 18. Stability comparison of models.

discriminator loss across 1000 training epochs. The loss function used to train the generator is a weighted sum of four different losses which composed of perceptual loss that utilize VGG features, edge loss to enhance structural clarity, adversarial loss to ensure realism, and total variation loss to reduce noise and promote smoothness. These losses are used together with weights which are chosen and tuned.

By using binary cross-entropy loss, the discriminator is trained so that it can distinguish correctly the real and generated images. As can be observed in the figure, the generator loss decreases stably and the discriminator loss converges without instability. The consistent reduction and balance between the two curves confirm robust and stable adversarial training. This supports the effectiveness of the loss design in guiding the generator to generate high-quality perceptually accurate results.

Figure 18 presents a comparison of stability analysis using the standard deviation of PSNR values. Both EDSR and RDN show the lowest value (0.30) whereas the MF-GAN is also competitive with 0.33, which is better than ESRGAN (0.37) and SRGAN (0.47) respectively that show higher variability due to adversarial training dynamics.

The performance-efficiency tradeoff, is illustrated in Figure 19, which presents peak PSNR as a function of computational cost (GFLOPs). The MF-GAN model shows an optimal balance between image quality and efficiency.

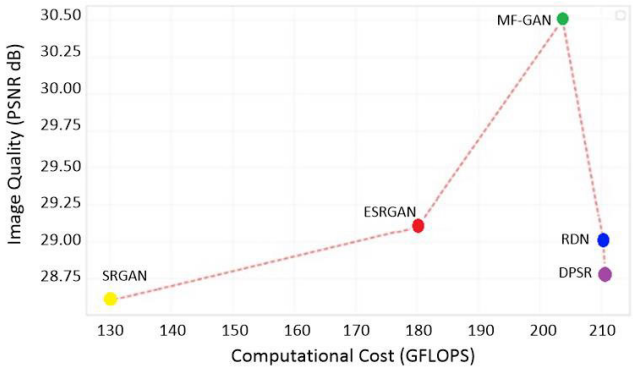

FIGURE 19. Performance efficiency analysis across GAN-based models.

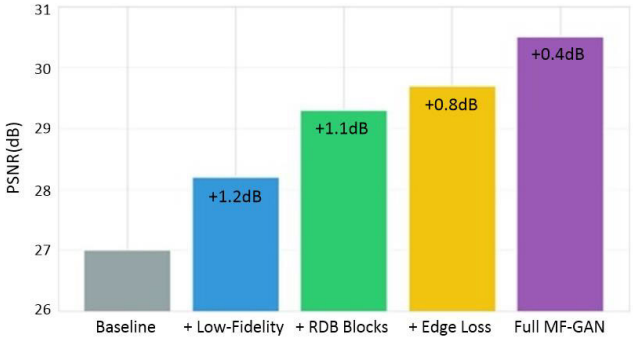

FIGURE 20. Performance analysis across GAN models.

Figure 20 visualizes the component-wise contribution to performance, revealing the incremental gains from low-fidelity structure, residual blocks, and edge loss. This form of ablation analysis illustrates the improvement in the performance contributed by each architectural.

C. STATISTICAL COMPARISON WITH STATE-OF-THE-ART MODELS

To further demonstrate the effectiveness of the proposed MF-GAN model, a detailed statistical comparison has performed with widely recognized state-of-the-art (SOTA) models, such as EDSR, RDN, SRGAN, and ESRGAN. The performance was verified with 4x scaling on three datasets (Set5, Set14, and DIV2K). All models were evaluated by both distortion-driven metrics (PSNR and SSIM) and perception-driven metrics (LPIPS), to achieve a fair comparison between fidelity and visual realism [36], [37].

The Table 11 show the average and standard deviation (±std) of the above metrics over five experiments per model on both statistical significance and reproducibility.

As shown in Table 11, EDSR and RDN achieve the highest PSNR as well as SSIM values, which is expected due to their MSE-focused optimization. They perform well in pixel level fidelity, but they are not capable to recover realistic texture restoration, as indicated by their higher LPIPS values.

On the contrary, MF-GAN has better perceptual performance, achieving the lowest LPIPS values over the three datasets, which indicates that is more capable of generating visually reasonable images. Although MF-GAN's PSNR is

TABLE 11. GAN model quantitative analysis.

Model	Dataset	PSNR (Mean±Std)	SSIM (Mean± Std)	LPIPS
EDSR	Set5	31.4 ± 0.3	0.93 ± 0.01	0.18± 0.02
RDN	Set5	31.2 ± 0.3	0.92 ± 0.01	0.19± 0.02
SRGAN	Set5	29.0 ± 0.5	0.89 ± 0.02	0.15± 0.02
ESRGAN	Set5	28.7 ± 0.3	0.83 ± 0.02	0.13± 0.01
MF-GAN	Set5	30.8 ± 0.3	0.92 ± 0.01	0.11± 0.01
EDSR	Set14	29.1 ± 0.3	0.88 ± 0.01	0.21± 0.02
RDN	Set14	29.0 ± 0.3	0.90 ± 0.01	0.19± 0.02
SRGAN	Set14	27.8 ± 0.5	0.86 ± 0.02	0.18± 0.01
ESRGAN	Set14	27.5 ± 0.5	0.87 ± 0.02	0.16±0.02
MF-GAN	Set14	29.5 ± 0.4	0.89 ± 0.01	0.13± 0.01
EDSR	DIV2K	30.2 ± 0.3	0.88 ± 0.01	0.22±0.02
RDN	DIV2K	29.6± 0.3	0.89 ± 0.01	0.21±0.02
SRGAN	DIV2K	28.1 ± 0.4	0.84 ± 0.03	0.19±0.03
ESRGAN	DIV2K	28.9 ± 0.3	0.86 ± 0.02	0.16± 0.02
MF-GAN	DIV2K	29.8 ± 0.3	0.88 ± 0.01	0.15± 0.01

slightly lower compared to EDSR and RDN (which is due to its adversarial training nature), its competitive PSNR are still achieved 30.8 $\pm$ 0.3 values on Set5, which is visually close to RDN's 31.2 $\pm$ 0.3.

A small difference in average PSNR and SSIM can be observed across repetition, which is natural in the GAN training process. Remarkably, the perceptual metrics of MF-GAN achieve consistent performance in both cases, demonstrating its stable generalization and visual realism.

These results represent the expected trade-off between distortion-based and perception-based objectives in GAN's frameworks. The proposed MF-GAN can address both structural integrity and perceptive reality through low and high-fidelity models, and represent an effective way to solve the real-world enhancement tasks. The evaluation is inspired by recent research into GAN where the best practice is to include the both peak performance and mean performance since both represent the full potential of the model as well as the average performance which represents its stability. The fact that MF-GAN achieves the SOTA visual quality marks its great practical efficiency, especially for tasks that value perceptual realism more than pixel level fidelity. In general, the proposed model leads to a statistically better performance than existing GAN-based models (SRGAN and ESRGAN) and remains a competitive solution when compared to the deterministic CNN-based methods.

VI. CONCLUSION

This study evaluated and compared the performance of MF-GAN for image enhancement with both current and conventional methods, including ESRGAN, RDN, DPSR, SRGAN, EDSR, Conditional GAN, as well as gamma correction and histogram equalization. With a PSNR of 31.4 dB and an SSIM of 0.92, MF-GAN achieves the highest image quality metrics, demonstrating greater structural integrity and better feature retention. ESRGAN and RDN also have showed strong results, however, in overall quality MF-GAN has higher quality, as well as it is capable of preserving fine details and textures that many other models frequently struggle with. In contrast, traditional methods performed worse with lower SSIM, PSNR, and perceptual quality values, indicating poorer structural fidelity and visual appeal. MF-GAN reduces computational redundancy, which has 3.1% less FLOPs than RDN and 3.3% less FLOPs than DPSR, while increasing PSNR by 1.4 dB over ESRGAN. MF-GAN's inference time of 1.32 seconds is slightly faster than the inference time of DPSR (1.34 seconds) by 1.49% but marginally slower than RDN (1.30 seconds) by 1.54%. This performance shows how traditional methods seem to fail in capturing complex details and textures. MF-GAN performance is due to its novel dual-generator design, which combines low-fidelity and high-fidelity models to gradually refine images. By including RDBs and a composite loss function, which includes perceptual, adversarial, and edge preservation components, MF-GAN effectively preserves textures and features that are often lost by other models.

It is ideal choice for high-resolution applications, which require preserving visual quality and structure, because its ability to preserve fine details. However, there is a restriction that only natural image datasets (Set5, Set14, DIV2K) have been tested on MF-GAN, and so far has not been tested on cross-domain such as medical images, satellite photos, or extreme low-light conditions. Adapting the model to such domains may require domain-specific tuning. Moreover, MF-GAN provides significant advantages for image enhancement, however, the dual-generator architecture causes a certain computational overhead which challenge its real-time use on mobile or embedded systems. Future work will explore the architectural simplifications and model compression solutions to further simplify MF-GAN and domain adaptation techniques to extend MF-GAN's application to various visual environments beyond natural scenes. This study shows the potential for advancing image processing by demonstrating the high structural similarity and image quality provided by the MF-GAN. The proposed method not only overcomes the drawbacks of current methods but also can serve as the stage for future developments in multi-fidelity image enhancement models.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[2] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[3] E. Lee and T. G. Whang-Bo, "Performance comparison of enhanced deep residual networks for single-image super-resolution model with change in loss function," in *Proc. Joint Int. Conf. Digit. Arts, Media Technol. ECTI northern Sect. Conf. Electr., Electron., Comput. Telecommun. Eng. (ECTI DAMT NCON)*, Phuket, Thailand, Mar. 2023, pp. 374–377, doi: 10.1109/ectidamtncon57770.2023.10139705.

[4] Z. Lin, S. Li, Y. Jiang, J. Wang, and Q. Luo, "Feedback multi-scale residual dense network for image super-resolution," *Signal Process., Image Commun.*, vol. 107, Sep. 2022, Art. no. 116760, doi: 10.1016/j.image.2022.116760.
[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, Jun. 2015.
[6] J. Liang et al., "A survey on evolutionary constrained multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 27, no. 2, pp. 201–221, Apr. 2023, doi: 10.1109/TEVC.2022.3155533.
[7] S.-W. Park, J.-S. Ko, J.-H. Huh, and J.-C. Kim, "Review on generative adversarial networks: Focusing on computer vision and its applications," *Electronics*, vol. 10, p. 1216, 2021, doi: 10.3390/electronics10101216.
[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[9] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 105–114.
[10] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.
[11] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.
[12] Z. Lu et al., "Multiobjective evolutionary design of deep convolutional neural networks for image classification," *IEEE Trans. Evol. Comput.*, vol. 25, no. 2, pp. 277–291, Apr. 2021, doi: 10.1109/TEVC.2020.3024708.
[13] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 694–711.
[14] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.
[15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
[16] X. Wang, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. Comput. Vis. ECCV Workshops (ECCV)*, in Lecture Notes in Computer Science, vol. 11133, L. Leal-Taixé, S. Roth, Eds., Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-11021-5_5.
[17] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D, Nonlinear Phenomena*, vol. 60, pp. 259–268, Nov. 1992.
[18] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th Int. Conf. Mach. Learn. (ICML)*, 2013, pp. 1–16.
[19] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Honolulu, HI, USA, Jul. 2017, pp. 1132–1140, doi: 10.1109/CVPRW.2017.151.
[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[21] S. S. Mohammed and H. G. Clarke, "Conditional image-to-image translation generative adversarial network (cGAN) for fabric defect data augmentation," *Neural Comput. Appl.*, vol. 36, no. 32, pp. 20231–20244, Nov. 2024, doi: 10.1007/s00521-024-10179-1.
[22] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 2472–2481, doi: 10.1109/CVPR.2018.00262.
[23] M. Krichen, "Generative adversarial networks," in *Proc. 14th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2023, pp. 1–7, doi: 10.1109/icccnt56998.2023.10306417.
[24] P. Zhou, D. Wu, X. Wang, C. Yan, and D. Zhou, "Non-uniform background noise suppression method based on improved total variation model for wide-field imaging system," *IEEE Access*, vol. 12, pp. 28141–28154, 2024, doi: 10.1109/ACCESS.2024.3365489.
[25] J. Xu, Y. Chae, B. Stenger, and A. Datta, "Dense ByNet: Residual dense network for image super resolution," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Athens, Greece, 2018, pp. 71–75, doi: 10.1109/ICIP.2018.8451696.
[26] S. S. A. B. Aldin, M. Aykaç, and N. B. Aldin, "Quad-color image encryption based on chaos and Fibonacci Q-matrix," *Multimedia Tools Appl.*, vol. 83, no. 3, pp. 7827–7846, Jan. 2024.
[27] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020, pp. 8107–8116, doi: 10.1109/CVPR42600.2020.00813.
[28] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proc. 7th Int. Conf. Learn. Represent.*, New Orleans, May 2019, p. 935. [Online]. Available: https://dblp.org/rec/conf/iclr/LoshchilovH19.html
[29] L. Yang and G. E. Karniadakis, "Potential flow generator with L_2 optimal transport regularity for generative models," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 528–538, Feb. 2022, doi: 10.1109/TNNLS.2020.3028042.
[30] S. S. Aldin, N. B. Aldin, and M. Aykaç, "Enhanced image classification using edge CNN (E-CNN)," *Vis. Comput.*, vol. 40, no. 1, pp. 319–332, Jan. 2024.
[31] İ. Babaoğlu, S. Kahveci, and A. Kılıç, "Enhanced pyramidal residual networks for single image super-resolution," *Neural Comput. & Appl.*, vol. 36, pp. 11563–11577, 2024, doi: 10.1007/s00521-024-09702-1.
[32] X. Min et al., "Perceptual video quality assessment: A survey," *Sci. China Inf. Sci.*, vol. 67, 2024, Art. no. 211301, doi: 10.1007/s11432-024-4133-3.
[33] X. Zhang, W. Lin, and Q. Huang, "Fine-grained image quality assessment: A revisit and further thinking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 5, pp. 2746–2759, May 2022, doi: 10.1109/TCSVT.2021.3096528.
[34] X. Wang, L. Sun, A. Chehri, and Y. Song, "A review of GAN-based super-resolution reconstruction for optical remote sensing images," *Remote Sens.*, vol. 15, no. 20, p. 5062, Oct. 2023, doi: 10.3390/rs15205062.
[35] M. Cheon, T. Vigier, L. Krasula, J. Lee, P. Le Callet, and J.-S. Lee, "Ambiguity of objective image quality metrics: A new methodology for performance evaluation," *Signal Process., Image Commun.*, vol. 93, 2021, Art. no. 116150, doi: 10.1016/j.image.2021.116150 .
[36] J. Yang, M. Lyu, Z. Qi, and Y. Shi, "Deep learning based image quality assessment: A survey," *Proc. Comput. Sci.*, vol. 221, pp. 1000–1005, 2023, doi: 10.1016/j.procs.2023.08.080.
[37] M. C. Neves, J. Filgueiras, Z. Kokkinogenis, M. C. F. Silva, J. B. L. M. Campos, and L. P. Reis, "Enhancing experimental image quality in two-phase bubbly systems with super-resolution using generative adversarial networks," *Int. J. Multiphase Flow*, vol. 180, Nov. 2024, Art. no. 104952.
[38] E. Pajouheshgar, T. Zhang, and S. Süsstrunk, "Optimizing latent space directions for Gan-based local image editing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 1740–1744.
[39] S. Perez, S. Maddu, I. F. Sbalzarini, and P. Poncet, "Adaptive weighting of Bayesian physics informed neural networks for multitask and multiscale forward and inverse problems," *J. Comput. Phys.*, vol. 491, Oct. 2023, Art. no. 112342.
[40] X. Qu, Z. Liu, C. Q. Wu, A. Hou, X. Yin, and Z. Chen, "MFGAN: Multimodal fusion for industrial anomaly detection using attention-based autoencoder and generative adversarial network," *Sensors*, vol. 24, p. 637, Jan. 2024, doi: 10.3390/s24020637.
[41] P. N. Srinivasu, G. L. A. Kumari, D. J. Kumari, P. Barsocchi, and A. K. Bhoi, "SegAN for recognition of caries from 2D-panoramic X-ray images," *IEEE Access*, vol. 13, pp. 100419–100432, 2025, doi: 10.1109/ACCESS.2025.3576914.
[42] B. Xu, D. Zhou, and W. Li, "Image enhancement algorithm based on GAN neural network," *IEEE Access*, vol. 10, pp. 36766–36777, 2022, doi: 10.1109/ACCESS.2022.3163241.

NOOR BAHA ALDIN received the B.S. degree (Hons.) in computer engineering from the University of Technology, Baghdad, Iraq, in 2008, and the M.S. and Ph.D. degrees in electrical and electronics engineering from Gaziantep University, Gaziantep, Türkiye, in 2012 and 2017, respectively. She was with the Electrical and Electronics Engineering Department, Gaziantep University, for around five years. She is currently a full-time Assistant Professor with the Electrical and Electronics Engineering Department, Hasan Kalyoncu University, Gaziantep. She has teaching experience of around six years. Her research interests include machine learning, image and video processing, wireless communications, and networking.