UNSUPERVISED MULTI-SCALE IMAGE ENHANCEMENT USING GENERATIVE DEEP LEARNING APPROACH

Preeti Sharma[1],[*], Manoj Kumar[2,3,4,5,*], and Hitesh Kumar Sharma[1]

[1] School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, 248007 India .
[2]School of Computer Science, FEIS, University of Wollongong in Dubai, Dubai Knowledge Park, Dubai, UAE
[3]Research Cluster Head, Network and Cyber Security, UOWD, Dubai
[4]MEU Research Unit, Middle East University, Amman, 11831, Jordan
[5]Research Fellow, INTI International University, Malaysia

ABSTRACT

To produce super-resolution images, it is essential to eliminate the noise elements and give a clear noise-free output. To achieve this purpose multiscale image representation is found to be effective in many ways for its accuracy of correct feature extraction capacity. This denoising approach is integrated as a chosen enhancement tool in the form of an ensemble GAN model, and accordingly, the generator-discriminator training concept is transformed to adopt the approach as per the desired demands. In this research, a multiscale image approach is implemented using an ensemble GAN model with hybrid discriminator architecture. No one form of noise is "ideal" to eliminate while denoising with the proposed model. Instead, based on the properties of the data and the noise inherent in it, the proposed ensemble GAN can handle various sorts of noise.The technique optimises training through simultaneous generator and discriminator model updates, improving output quality, by using the least loss value for discriminator selection. Inception Score (IS) and Fréchet Inception Distance (FID) evaluations show that it outperforms pixel-based denoising, with an amazing accuracy of 99.91%.

KEYWORDS

GAN, multiscale image representation, ensemble GAN, pixel based denoising, Multiscale denoising.

1. INTRODUCTION

The deep learning research division is becoming increasingly detailed in Generative Adversarial Networks (GANs) in a multidisciplinary manner. GANs are widely used to generate high-quality playback images used in several fields, such as forensics, medical diagnostics, architecture, and filmmaking. GAN versions are currently used to create images from text data, create movies from still images, increase image resolution, and manipulate images. Its uses range from detecting anomalies to improving chess games [1]. Image noise is a common problem and GAN works to refine and restore low-quality or degraded images. Using the contingency distribution of a collection of media data GAN generates new and accurate visual results to improve denoising as shown in Figure 1.

David C. Wyld et al. (Eds): ICCSEA, NLAI, SCAI, CSIA, IBCOM, SEMIT, NECO, SPPR, MLDS -2023
pp. 193 – 205, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.132415

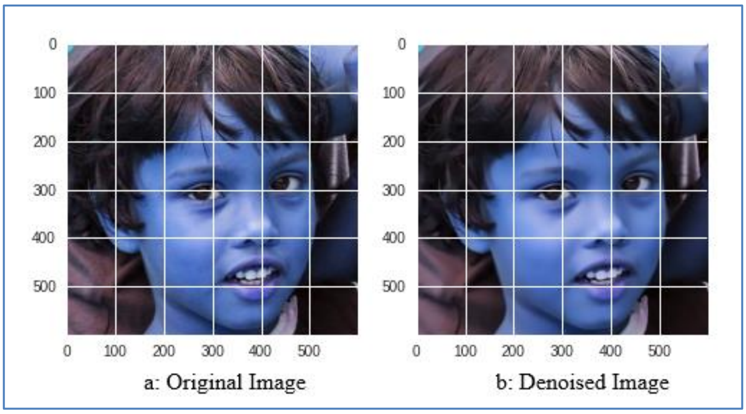

Figure 1: Denoising of Coloured Image using Ensemble GAN Model

It has been observed that a captured image of a real scene or object can be degraded by inaccurate capture and optical factors such as improper depth of field, poor focus, camera shake, object movement, short exposure, and poor optical quality. Tasks involving image identification or classification can be impeded by noisy visuals with imperfections[2]. Noise often comes out in images as detached pixels or blocks of pixels, causing heavy visual effects. In general, the noise signal is inappropriate for the test object. It emerges as impractical information that disturbs the picture's visible information and is a significant barrier to image processing. Therefore, recovering the original signal and preserving features from noisy images is the goal of the extremely promising work of image denoising in graphics [3].

The current advancement in picture denoising has been focused on deep denoiser, also known as the deep network for denoising. Unsupervised deep denoisers that just call chaotic noisy photos without training data have gained popularity over the last few years. Relaxing the need of supervised learning on training samples has recently attracted growing interest. The training of DNNs using a noisy picture dataset without pair-wise correspondence or even just the input noisy image itself has been the subject of more recent research on unsupervised deep denoisers. These techniques fall into one of two types [4]: -

- Data Augmentation Method: - When training a DNN to map a noisy image to itself, Noise2Void and Noise2Self use the blind-spot strategy to avoid overfitting (convergence to identity map), while Noiser2Noise and Noise-as-Clean add additional noise to the original noisy image to produce image pairs that are then used to train the DNN.

- Denoising DNN Regularization: - By compensating the prediction's divergence, Stein's Unbiased Risk Estimator (SURE) regularizes the DNN. Early-stopping is used in the Deep Image Prior to prevent overfitting. To lessen the bias and variance of the prediction from the DNN trained on a single noisy picture, Self2Self introduces a dropout-based training/testing strategy.

As according to the latest studies, deep learning has revolutionized the fields of pattern detection and computer vision. k-means singular-value decomposition, Principal component analysis using local pixel grouping (LPGPCA), Block matching and 3D filtering (BM3D), and weighted nuclear norm minimization (WNNM) are examples of conventional denoising techniques. These techniques are made to reduce noise depending on the characteristics of both the pictures and the noise. To map from noisy photos to clean images, learning-based techniques like a denoising convolutional neural network (Dn_CNN) frequently employ paired-image datasets[5]. Among these techniques, convolutional neural networks (CNNs) have been proven to perform well in the disciplines of visual identification, object identification, lossy compression, visual super-resolution, and visual noise removal, amongst many other digital image challenges.The prevalence of Generative Adversarial Networks (GANs) in image enhancement expanded, following their success with accurate image reconstruction results. GAN with help of its adversarial training approach work initiates with noisy image construction and by optimizing loss value of both generator and discriminator, it generated noise free super resolution images. The mathematical formulation of GAN is listed in equation 1.

$$\boldsymbol{GminDmax\,V(D,G)} = \boldsymbol{Ex} \sim \boldsymbol{pdata}(\boldsymbol{x})[\boldsymbol{logD}(\boldsymbol{x})] + \boldsymbol{Ez} \sim \boldsymbol{pz}(\boldsymbol{z})\left[\boldsymbol{log}\left(\mathbf{1} - \boldsymbol{D}\big(\boldsymbol{G}(\boldsymbol{z})\big)\right)\right] \qquad \mathbf{(1)}$$

Equation 1: Mathematical formula of GAN [6]

Where; G stands for Generator, D stands for Discriminator, Pdata(x) = real-world data distribution, P(z) = generator distribution, x = Pdata sample (x), z = a sample taken from P (z), D(x) denotes a discriminator network, G(z) denotes the generator network.

GAN basically uses multiscale representations technique for analysing and modelling tasks which has huge importance in imaging utilization. In contrast to traditional methods, multiscale representations of convolutional neural networks are primarily viewed as feature pyramids. Feature pyramids can leverage further through convolution and down sampling processes, visual background data is converted across local to global perspective. Also, utilizing adversarial training, GANs might assistance overall image restoration to yield HR picture patterns that are greater realistic and precise. However, LR picture characteristics and subsequent up sampling techniques are constrained by GAN-based ultra-composite restorations. In contrast, the LR imagery' limited dimensions often cause excessive noise to exist in the reformed output. Therefore, optimizing the texture receptive field of the network is difficult and cannot fully exploit multi-scale contextual information [7].

This research work utilizes multi-scale representations of images using ensemble GAN approach as an effective noise reduction tool. The model built using one generator coupled with three discriminators architecture. The model utilises the loss value of each discriminator for starting 50 epochs and generates a raw image including noise. Based on least loss value, the discriminator is finalised for execution with generator for remaining epochs. The image generated will then turned into super resolution image by fine tuning in subsequent iterations of the model.

The architecture work like a multi-scale denoiser to improve image output quality and ensure noise-free reproduction of images. Also, mode training problems such as mode collapse and non-convergence of results can also be eliminated using this ensemble approach. Overall, this model is said to offer the advantage of better image restoration while minimizing the potential for distortion and error. The model is evaluated using the Inception Score (IS) and Fréchet Inception Distance (FID) parameters. It outperforms to other conventional approaches showing accuracy score of 99.91%.

Rest of paper is structured with subsequent sections. Sections II and III presented the background and related research of the domain. Section IV defines the proposed methodology including the details about the dataset and defined architecture. Section V demonstrates the experimental and results section. It demonstrates various result graphs along with the comparative study of the proposed model with existing models Deep Convolutional GAN (DCGAN) and Conditional GAN (CGAN). Finally, section VI concludes the research paper.

2. BACKGROUND STUDY

Images can contain large amounts of data and instruction, so image deputation techniques have applications in many fields such as medicine, conveying, military, aerospace, and communications. It is ingrained in our lives and is attached from each of us. However, image noise can occur during image acquisition, storage, and processing, disrupting the instruction provided by the image and reducing image clarity. Image noise generally includes salt and pepper noise, gamma noise, and Gaussian noise. To solve this problem, high-resolution image denoising and image reconstruction (SR) techniques were researched to restore high-quality, high-resolution images[8]. Formally, the solution to the denoising limitation is based on three basic components[9]:

A signal model is followed by a noise model and a signal fidelity assessment (majorly recognized as the objective function). The concept of In Mathematical model, visual degeneration can be characterized as x = y + n, where x represents the degraded version of the original image y and n represents the added noise, also known as Additive White Gaussian Noise (AWGN), as illustrated in Figure 3 below.

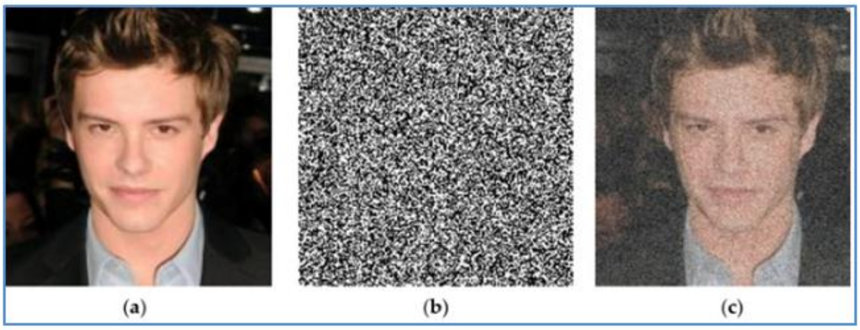

Figure 2: (a) Exposes an unedited or true visual (y), (b) Showcases the AWGN visual (n), and (c) reveals the $(x = y + n)$ resultant image [9].

Tavakoli *et al*. [10] provides an approach approach that is effective for picture denoising, according to simulation findings. The original picture is rebuilt using the observation vector and existing recovery methods like L1 minimization after an unknown noisy image of interest is detected (sensed) using a small number of linear functions in random projection. Rajni et al.[11] employ a novel methodology called Optimal Wavelet Basis (OWB). OWB uses Shannon entropy to choose the optimal tree from a noisy picture by applying multilayer Wavelet Packet (WP) decomposition.An innovative technique for de-noising MR (Magnetic Resonance) pictures with Rician noise has been developed by Choudhary et al[12]. This approach improves the Morphological Mean Filter (MMF).

Image denoising methods focus on recovering a visual that has been denoised from a laterally noisy image (x) by removing or contracting the noise (n). With these techniques, accompanying white noise and Gaussian noise are denoised using multiresolution or multiscale analysis.A simple way to denoise an image is to separate the image into equivalent low and frequency-heavy content. Then, this decline can be modified iteratively to the low sub bands. Low sub bands comprise to create high frequency sub band which combines to generate a multiscale representation. It contains the identical data as the reference image without noise as shown in Figure 2.

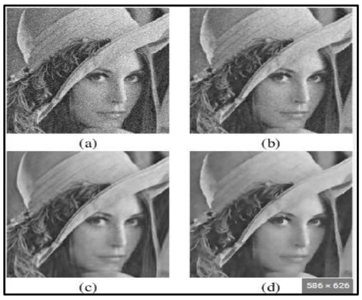

Figure 3: Multiscale Denoising Layers up to Final Output[13]

Image denoising and super-resolution (SR) are two important stages of image processing and are usually studied separately. But, for visual improvement, image noise reduction should be combined with image super-resolution process. Several unsupervised image enhancement methods based on deep learning have been recently developed to eliminate the dependence on paired data (x, y). These methods use Generative Adversarial Networks (GANs) to approximate the handling of generated images to that of target images without pairwise learning. As an example, severally image-to-image conversion or conversion models such as Cycle GAN can be applied for image enhancement. Also, GAN-based models were also worked to address the task of lighting intensification. These unsupervised models can produce better lighting and colour images. However, in real-world low-light image enhancement tasks are bounded by few constraints listed as:

- Enhanced image contrast and lighting may sometimes result in colour distortions and inconsistencies. Bright areas of dark images may be overexposed.
- Due to unstable training of unsupervised models, contiguous regions may performance sharp colour or illumination discrepancies.
- Models with a single frame-to-frame mapping network, when applied to low-light, high-noise images, are primarily concerned with illumination enhancement, but consistently fail at noise reduction.

As a solution, multi-scale image display is integrated into these models to ensure improved image quality and noise reduction.

3. RELATED WORK

This section consists of a selection of current published references that highlight the importance of noise reduction tools for GAN image regeneration applications. The selection of literature is systematic, following two main aspects (a) Concerns about noise reduction in GAN image SR models. (b) Using multi-scale image representations in GAN models as a solution for image noise devaluation.Yan *et al.*[14] conferred his self-consistent GAN (SCGAN), which allows unchecked noise and can dynamically extract noise maps from noisy images modelling. It received three new self-consistent compulsion that complement each other.

- A clean input should result in a zero response from a noise standard.
- Once presented with a pure noise input, a noise model ought to produce the same result.
- Once the pure noise map is coupled to the clean image, the noisy model also has to extract the relevant to the pattern.

Three major image processing tasks, including blind image denoising, rain streak removal, and noise image super-resolution are achieved using simplicity and effectiveness of the GAN model. It also highlights the importance of noise reduction tools for GAN image regeneration applications focusing on two main aspects. (a) Concerns about noise reduction in GAN image SR models. (b) Using multi-scale image representations in GAN models as a solution for image noise devaluation.

Deep Convolutional Neural Networks (CNN)-based novel algorithm has been suggested by Lee *et al.* [15]. The technique is built on the Deep Convolution Network, which consists of several U-nets. Each U-net reduces noise of different intensities and is graded according to performance improvement. rough. Additionally, the combined CNN employs a comprehensive 3D convolution method. The potential standard might speed up end-to-end learning beyond pre- and post-processing thanks to such an architecture.

Li et al. [16] granted a revolutionary generative adversarial network (MSAt-GAN) based on deep attention mechanism and multiscale feature transmission blending. The proposed model was used for fusing visible and infrared imagery. In this study, rather than artificially fixing a single receptive field, he used three different receptive fields to capture multiscale and multilevel depth features of multimodality figure in three channels.

- Firstly, the crucial features of the origin form could be improved extracted from distinctive accessible fields and angles, as well as more adaptable and diverse feature representations, were also featured.
- Secondly, a multiscale fusion mechanism for deep attention was designed. Third, the chained operations to improve feature submission while achieving improved usage of preceding features, multi-level deep features in the encoder and deep features in the decoder are cascaded.
- Ultimately, a GAN with two discriminators was advanced on top of the network structure. As a result, the created image simultaneously preserves the intensity of the infrared image and the texture details of the visible image.

Wang *et al.*[17] To fulfil the image denoising goal, a Deep Residual Network based on Generative Adversarial (GAN) networks was proposed. First, a residual block-based generative adversarial network structure was created. The GAN network was then trained using an advanced loss function. The resulting image is very closely resembled by a well-designed loss function including its distinct counterpart (the ground truth) while enhancing colour and brightness details.

4. RESEARCH METHODOLOGY

4.1. Model Development

The architecture is based on multiple (3) discriminators coupled with single generator are included in the GAN ensemble model. The loss function used in this model architecture is resolved using a minimum distance criterion between the generated distribution and the actual allocation. Initially, the model utilizes the loss value of each discriminator for starting 50 epochs and generates a raw image including noise. Based on least loss value, one discriminator is finalized for execution with generator for remaining epochs which defines denoising process. The image generated will then be turned into super resolution image by fine tuning in subsequent iterations of the model.

By maximizing the objective function regarding the generator and reducing it with respect to the Discriminator, both modules are set against one another (it should be noted that in the regular procedure, we just minimize a loss function). With G and D, the loss function is denoted by the notation E(D,G). Then, E will be the sum of the number that D predicts as 1 for the false picture Xfake or D(z) plus the number that D predicts as 0 for the true image Xreal. For optimization of objective function listed in equation 2 we try to achieve GminDmax E(D,G) as mentioned below,

$$GminDmax\ E(D,G)\ = 1/2[\ Ex \sim px(1 - D(x))] + 1/2[Ez{\sim}p(z)D(G(z))] \quad (2)$$

This novel approach for the GAN is designed to address the following issues:

(a) an optimized discriminator D (better approximating Max E (D, G)).
(b) a D better suited to the generator's G capabilities.
(c) overcoming common flaws of other GANs, such as a collapse of the global framework and inconsistency in local details; and
(d) integrated with multiscale denoising approach for super-resolution output using ensemble technique.

Two types of pair training illustration are used in this study. That is, the noisy input image xiNi=1X and t target image should be YiNi=1Y is set. The fabrication of a noise-free image V(x) that resembled the already-existing clean target image was used to qualify a denoising network V. A set of three discriminators D1, D2, and D3 are trained concurrently to distinguish the faked actual clear photographs that are noise-free. By reducing adversarial loss and attempting to trick the discriminative network, the denoiser learns to convert noisy regions into clean genuine regions. An encoding network E n, a residual block layer R, and a decoding network De make up the noise reduction network. The encoder is made up of several convolutional downsampling developments that reconstruct a noisy image into a feature domain (x).

$$V(x) = De(R(En(x))) \quad (3)$$

These feature domains E n(x) are then inserted into the remaining block. The decoder network De receives the residual block R(En(x)output)'s feature map as an input. At this phase, the rectified feature map is decoded into a fabricated, clean image using a series of upsampled, transposed convolutional layers (x).

4.2. Generator and Discriminator Network

Indistinguishable pictures are produced using a generator network so that the discriminator cannot identify them apart from actual images. Images that are equivalent to real high-definition image labels are created using noisy picture inputs, but they must outperform the labels in terms of details and colours. An eight-layer convolutional architecture is employed for this network. It is to be taken care that all the attributes of our input image must be kept and improved upon, and the form of the images was unaltered throughout the process [13]. Research demonstrates that the network can keep more details as compared to conventional architectures. The discriminator Network includes a stack of three different discriminators. These uses different deconvolutional networks.

Simple convolutional layers connect to a complete link layer in each discriminator's network configuration, and the confidence score is then normalized using the sigmoid function. The best voting classifier approach is used to ensemble the three discriminators to the generator. If the network has been properly trained, it will give a score of close to 1 for the label picture that exists and a score of 0 for the fake image, demonstrating that it has a good capacity for discrimination and can successfully distinguish between real and fake images. Each discriminator's loss value determines which discriminator will be used at each epoch to construct the GAN model. In our experiment, the true labels are created from generated noised images, whereas the labels for the photograph are clear images. Figure 4 depicts the suggested GAN model's architecture.

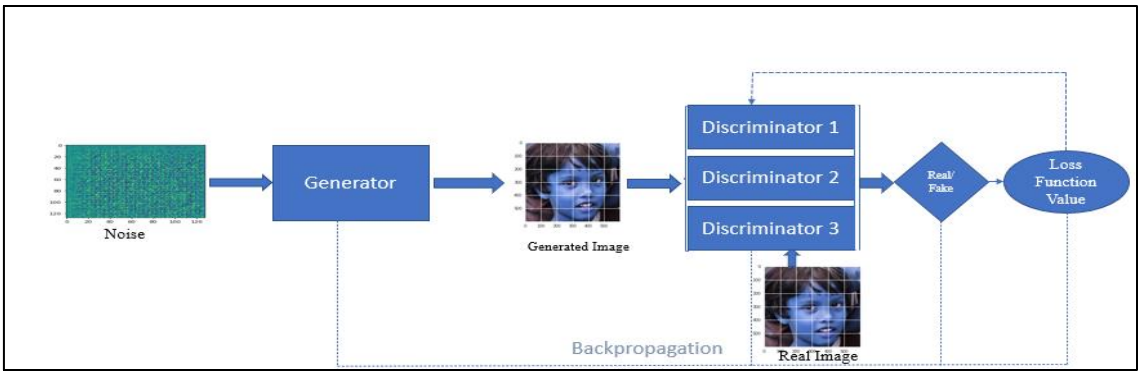

Figure 4: GAN Model with Multiscale Denoiser using Ensemble Approach.

4.3. Training

The Ensemble GAN model typically features non-monotonic loss functions and are very hypersensitive to parameters. It gets trained to do a variety of tasks by altering the input. So, pre-training the generator and discriminator is typically required. To make up-scaled images for super-resolution, for instance, we feed in downscaled photos and let the generator network create them. We use noisy pictures for denoising and gaussian kernel-blurred images for deconvolution. Batch normalization was applied except for the final layer, to each of the convolutional layers and residual blocks for both the generative and discriminator networks, to avoid gradients from contracting or expanding.

Denoising algorithm of the proposed model:

Algorithm: GAN Model with Multiscale Denoiser using Ensemble Approach.

Input: Data Samples and Noise Samples of Face Images.
Step 1: Pre-training the GAN generator (Gn1) by fixing reconstruction parameters.
Step 2: GAN fake denoised generation (Gn2) through adversarial training with ensemble discriminator feedback.
Step 3: Deep learning Ensemble Detectors (ED) applied to verify the GAN fake generator effect (Gn1+Gn2).
Step 4: Optimize and update each of discriminator in the ensemble using loss function value generated.
Step 5: Minimize generator loss function and update.
Step 6: Fake and Real Image Classification Analysis output.

4.4. Dataset

Google Facial Expression Comparison (FEC) [18] is used for the purpose of training and testing the proposed model. This dataset of faces was collected from Flickr.com to create software for facial expression detection and search. The dataset consists of 87,517 verified photos that were obtained from Flickr and utilised for biometric analysis. Exposing.ai discovered that the 87,517 distinct photographs belonged to 45,382 distinct Flickr customer accounts. The FEC dataset is positioned within the context of bigger facial recognition datasets like MS-Celeb-1M and Mega Face datasets. The dataset that consists of large scale and heterogeneous facial image information is pre-processed with necessary refinements to improve model training and obtain better accuracy with minimum time consumption. These pre-processing tools apply cropping for faster computation of loss function and identify suitable normalisation. On the sampled data set, the training and test samples are divided 7:3 respectively.

4.5. Evaluation

In this work, Accuracy, Loss values, Fréchet Inception Distance (FID) and Inception Score (IS) that are suggested as ad-hoc metrics to evaluate the performance of model. These metrics are indicators of the visual quality of created images.

The Inception Score devised by Salimans et al. accurately depicts the photographs' quality and diversity. The IS measures the estimated difference between the spread of class labels used to train the external network and the dispersion of class predictions for data from the GAN. It is calculated using the Kullback-Leibler (KL) divergence, a statistics formula which measure of how similar or dissimilar two probability. Our distributions differ when KL divergence is substantial. That is, it is high when the whole set of created photos has a wide variety of labels, and each generated image has a unique label. The KL divergence's exponential is used to get the final score, and then its average across all of our photos is used to calculate the IS.

The Fréchet Inception Distance (FID) devised by Heusel et al. extends IS is a statistic that determines how far apart feature vectors determined for actual and artificially created pictures are from one another. The score enumerates the statistics on computer vision aspects of the original pictures that were derived utilising the Inception v3 model to classify images, and it compares the two groups. A perfect score of 0.0 indicates that the two sets of photographs are identical, whereas lower values show that the two groups of photos are statistically more alike or similar. It is used to assess the level of picture quality produced by generative adversarial networks, and lower scores have been found to be positively correlated with better images.

5. Experiments and Results

In contrast to conventional image processing techniques, the suggested multiscale Ensemble GAN approach enables us to employ a single architecture framework to accomplish many goals. To train the model, we just need to change the pre-processing stage and add new inputs. The generator is forced to create images that seem better due to rivalry between it and the discriminator. The model may use learnt features to create pictures from inputs that lack specific information since it can learn from large datasets. For instance, the suggested model can construct human faces that seem convincing even when given extremely low-resolution photos of human faces as input. The model performance on loss function in generating the noise free images using multiscale Ensemble GAN from images with noise as input is given below in Figure 5:

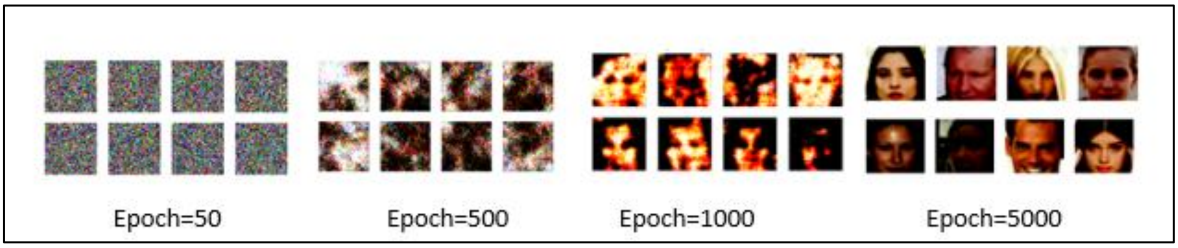

Figure 5: Experimental Output of Multiscale Denoiser in the Proposed GAN Model

The proposed multiscale ensemble GAN model is worked using multiscale denoising approach to get the best super-resolution image output as well as minimize the model training problems. The model is fitted well with reasonable modifications. It helps to eliminate model training difficulties and provide a simplified loss function to confirm optimal model output. Model performance is ensured with the help of 3 discriminator ensemble that is trained aligned with generator so that they can correctly locate denoising as is applied in the generator output. Alongside, the discriminator is chosen at each iteration based on its result that minimizes loss function. The model outcome is evaluated with accuracy of 99.91%, IS value1.4 and FID value of 412.

The graph below shows the difference level of real input and fake image generation as produced by the model presented in Figure 6 and Figure 7. Model achieves good optimized values of IS (1.4) and FID(412) shown in the graph below mentioned as Figure 8 and Figure 9 respectively. It is observed to outperform, achieving great accuracy of 99.91%.

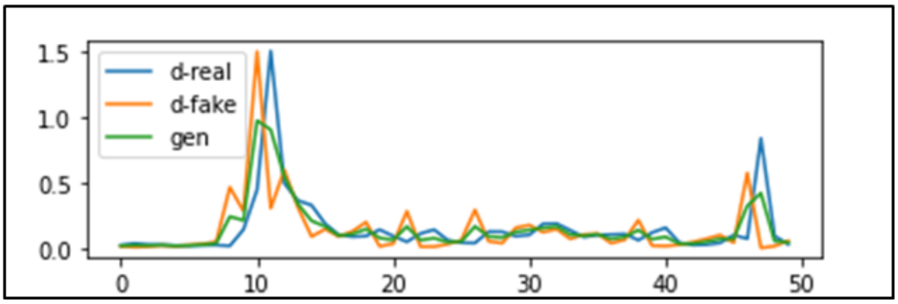

Figure 6: Line Plots of Multiscale Ensemble GAN model with improved denoising image generation graph.

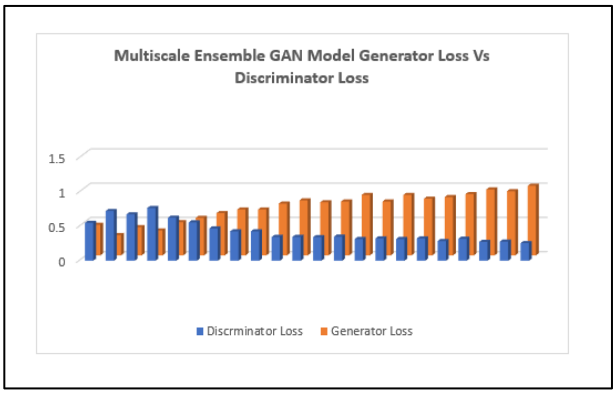

Figure 7: Line Plots of Multiscale Ensemble GAN Model Loss Function of generator and discriminator.

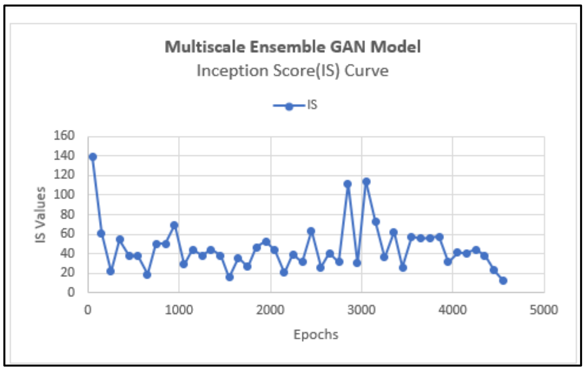

Figure 8: Line Plots of Multiscale Ensemble GAN Model Inception Score (IS).

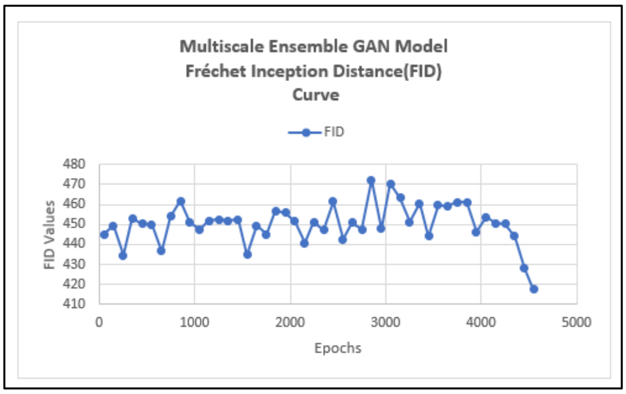

Figure 8: Line Plots of Multiscale Ensemble GAN Model Fréchet Inception Distance (FID).

The model's generalization and robustness are provided by the IS (1.4) and FID (412) scores' optimized values. The model performs excellently and achieves a respectable accuracy rating of 99.91%. On Basisof achieved scores it is observed that the architecture works like a multi-scale denoiser to improve image output quality and ensure noise-free reproduction of images. Also, mode training problems can also be eliminated using this ensemble approach. Performance studies suggest that model analyses the image with many typical facial features and detects facial details rather well. Additionally, the model can conduct deconvolution on human faces in a respectably acceptable manner. It is observed to be better than the other approaches as shown in Table 1 below.

Research	**Dataset**	**Accuracy Score**
Ours	Google Facial Expression Comparison (FEC)	99.91%
Wang et al. [19]	CIFAR10	99.5%
Benny et al. [20]	CIFAR10 / MIST	69.51% 98.90%

When compared to current models, our suggested model outperforms them. Our model achieves an accuracy of 99.91% using the Google Facial Expression Comparison (FEC). When applied to CIFAR10, Wang et al. attain an accuracy of 99.5%, which is surpassed by our model. Furthermore, when compared to Benny et al.'s models on the CIFAR10 and MIST datasets, which reach 69.51% and 98.90% accuracy, respectively, our model exceeds both in terms of accuracy, indicating its usefulness in accurately capturing and recognizing facial emotions. The findings demonstrate our suggested model's improved performance, establishing it as a strong challenger in the domain of facial expression comparison.

6. CONCLUSION

The presented research provides a strong technique to denoising and super-resolution via the combination of multiscale picture representation with an ensemble Generative Adversarial Network (GAN) model. Using a hybrid discriminator architecture designed for multiscale analysis, the model exhibits a surprising capacity to manage various types of noise inherent in the input data. The training approach, which includes simultaneous updates to the generator and discriminator models, optimises the learning technique. Furthermore, dynamically choosing the discriminator with the lowest loss value improves overall output quality. Inception Score (IS) and Fréchet Inception Distance (FID) evaluation measures highlight the performance of the proposed ensemble GAN model, with an excellent accuracy of 99.91%. This study not only enhances the state-of-the-art in denoising techniques, but also demonstrates the versatility of ensemble GANs for dealing with diverse forms of noise in multiscale picture representation, resulting in high-quality, noise-free images.

In terms of future work, there are still several restrictions on the suggested GAN model, despite its ability to provide realistically synthesised discrete data, continuous data, and even time series data to address the problems of fewer labels and imbalanced classifications. It has been difficult to analyse these created data and figure out how to apply them to current applications, such as medical research, which requires genuine data to corroborate.

REFERENCES

[1] Y. Intrator, G. Katz, and A. Shabtai, "Multi-Discriminator Generative Adversarial Networks".

[2] R. A. Khan, Y. Luo, and F. X. Wu, "Multi-scale GAN with residual image learning for removing heterogeneous blur," *IET Image Process.*, vol. 16, no. 9, pp. 2412–2431, 2022, doi: 10.1049/ipr2.12497.

[3] P. Gong, J. Liu, and S. Lv, "Image Denoising with GAN Based Model," *J. Inf. Hiding Priv. Prot.*, vol. 2, no. 4, pp. 155–163, 2020, doi: 10.32604/jihpp.2020.010453.

[4] T. Pang, H. Zheng, Y. Quan, and H. Ji, "Recorrupted-to-Recorrupted: Unsupervised Deep Learning for Image Denoising," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2043–2052, 2021, doi: 10.1109/CVPR46437.2021.00208.

[5] L. D. Tran, S. M. Nguyen, and M. Arai, "GAN-Based Noise Model for Denoising Real Images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12625 LNCS, pp. 560–572, 2021, doi: 10.1007/978-3-030-69538-5_34.

[6] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.

[7] C. Yuan, K. Deng, C. Li, X. Zhang, and Y. Li, "Improving Image Super-Resolution Based on Multiscale Generative Adversarial Networks," 2022.

[8] S. Li, P. Qian, X. Zhang, and A. Chen, "Reconstruction Technology of Multiscale-Fusion Images," vol. 2021, 2021.

[9] C. C. By, "Adversarial Gaussian Denoiser for Multiple-Level," 2021.

[10] [A. Tavakoli and A. Pourmohammad, "Image Denoising Based on Compressed Sensing," *Int. J. Comput. Theory Eng.*, vol. 4, no. 2, pp. 266–269, 2012, doi: 10.7763/ijcte.2012.v4.463.

[11] R. Rajni, “Image Denoising using Wavelet Packet Transform and Optimal Wavelet Basis Image Denoising using Wavelet Packet Transform and Optimal Wavelet Basis,” *Proceeding Natl. Conf. Comput. Commun. Electr. Syst. 2016*, no. November 2016, 2018.

[12] J. Choudhary and A. Choudhary, “Enhancement in Morphological Mean Filter for Image Denoising Using GLCM Algorithm,” *Int. J. Comput. Theory Eng.*, vol. 13, no. 4, pp. 134–137, 2021, doi: 10.7763/ijcte.2021.v13.1302.

[13] M. Tico, “MULTI-FRAME IMAGE DENOISING AND STABILIZATION,” no. Eusipco, pp. 1–4, 2008.

[14] H. Yan, X. Chen, V. Tan, W. Yang, and J. Wu, “Unsupervised Image Noise Modeling with Self-Consistent GAN”.

[15] S. Lee, M. Negishi, H. Urakubo, H. Kasai, and S. Ishii, “Mu-net : Multi-scale U-net for two-photon microscopy image denoising and restoration,” *Neural Networks*, vol. 125, pp. 92–103, 2020, doi: 10.1016/j.neunet.2020.01.026.

[16] J. Li, B. Li, Y. Jiang, and W. Cai, “MSAt-GAN : a generative adversarial network based on multi-scale and deep attention mechanism for infrared and visible light image fusion,” *Complex Intell. Syst.*, vol. 8, no. 6, pp. 4753–4781, 2022, doi: 10.1007/s40747-022-00722-9.

[17] Z. Wang, L. Wang, S. Duan, and Y. Li, “An Image Denoising Method Based on Deep Residual GAN,” *J. Phys. Conf. Ser.*, vol. 1550, no. 3, 2020, doi: 10.1088/1742-6596/1550/3/032127.

[18] R Vemulapalli, A Agarwala, “ A Compact Embedding for Facial Expression Similarity ”, CoRR, abs/1811.11283, 2018.

[19] Z. Wang, Q. She and T. Ward, "Generative Adversarial Networks in Computer Vision", (2022), *ACM Computing Surveys*, vol. 54, no. 2, pp. 1-38. Available: 10.1145/3439723.

[20] Y. Benny, T. Galanti, S. Benaim and L. Wolf, "Evaluation Metrics for Conditional Image Generation" (2021), I*nternational Journal of Computer Vision*, Vol. 129, no. 5, pp. 1712-1731. Available: 10.1007/s11263-020-01424-w.

© 2023 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license