Enhancing Image Quality Using Advanced Deep Learning Techniques

Sajja Radha Rania, Modala Venkata Ranga Satya Pallavi, Chikkudu Sri Vardhan, Mannem Naveen Kumar Reddy and Kalwa Eswar

Department of Advanced Computer Science and Engineering, Vignan's Foundation for Science, Technology and Research, Guntur, Andhra Pradesh, India

Keywords: Computer Vision, High Resolution Imaging, Generator-Discriminator Framework, Super Resolution, Adversarial Training, Perceptual Loss, Content Loss, Generative Adversarial Networks, Adversarial Loss, Convolutional Neural Networks.

Abstract: This work introduces a new approach for image super resolution using a specifically designed Generative Adversarial Network architecture with significant improvement in facial recognition. Our approach involves a two-stage training procedure where the discriminator is pre-trained on the whole dataset to strongly separate detailed features and minute differences between real images. After the discriminator builds a consistent baseline, the entire GAN model is trained with a reinforced generator including extra layers with the purpose of filtering noise and maintaining fine features. By enhancing the generator structure, the developed method not only generates high-quality images but also maintains important facial details, hence overcoming typical pitfalls in low resolution image re- construction. Experimental outcomes verify that this augmented GAN design successfully fills the gap between low resolution inputs and high-quality outputs, providing significant potential for applications in real life in face recognition and more.

1 INTRODUCTION

Super resolution of images is a crucial research area that aims to restore low resolution images into high resolution images, recovering much detail lost during image acquisition or compression. The major aim is to recover images that not only are clearer but also retain the authenticity of their original details. Traditional interpolation methods (Zhang et., al. 2018) are often at shortfalls in this regard, resulting in blurred or artifact-ridden output. Conversely, contemporary deep learning techniques, Convolutional Neural Networks, have shown immense success by using their layered structure to extract low-level textures and high-level semantic information. This development opens doors to more precise and realistic image reconstructions, which are vital for a variety of real-world applications.

The history of CNNs in image processing has been impressive because they can learn hierarchical representations from data. By piling multiple convolutional layers, such networks are well-positioned to pick up features on different scales and complexities ranging from basic edges and textures to higher- level patterns and contextual signals. Nevertheless, much as their performances are remarkable given these virtues, CNN- based strategies occasionally under-perform against some of the limitations Lin and Shum (2004) inherent in the super resolution processes. Our proprietary generative method is spearheaded by its aim to harness the maximum yield and enrichment of facial features out of low resolution images, where the requirement especially becomes a mission-critical operation in surveillance systems. Conventional super resolution techniques tend to have difficulty with noisy inputs, which can produce suboptimal or distorted outputs that do not capture critical facial details. In order to counter this problem, we have incorporated a denoising module directly into our GAN architecture. This adjustment is specifically designed to counteract the noise commonly found in surveillance video (M. Elad and M. Aharon, 2006), producing cleaner outputs that are more likely to retain facial features. The clean, denoised images generated by our model can then be used as better inputs for more sophisticated facial restoration

Rania, S. R., Pallavi, M. V. R. S., Vardhan, C. S., Reddy, M. N. K. and Eswar, K.
Enhancing Image Quality Using Advanced Deep Learning Techniques.
DOI: 10.5220/0013902200004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 3*, pages 579-586
ISBN: 978-989-758-777-1

methods, like GFPGAN (Xintao, et al. 2021), that further refine and sharpen facial features. This two-step, combined process thus successfully circumvents the drawback caused by noise to deliver better accuracy and reliability in face recognition in high-stakes use cases.

In addition, our GAN model is particularly altered in that it consists of extra layers in the generator, which are used to absorb a wider range of image characteristics and reduce noise. This denoising feature (M. Elad and M. Aharon, 2006), is especially useful in situations where images are contaminated by environmental noise like low-light or sensor noise. For instance, improving night vision surveillance footage or recovering corrupted photos is possible with our method, as the network successfully eliminates un- wanted noise while reconstructing missing details at the same time. The incorporation of such denoising capabilities guaran- tees that output images not only display better resolution but also preserve the sharpness and integrity required for important operations.

This paper is structured into the following sections: section II contains previous work in the image super resolution field. section III gives a clear outline of the generative adversial networks. section IV explains the methodology applied in the current paper to achieve the desired outcome. section V give experiments and results while Section VI summarizes the work and highlights potential areas for future exploration.

2 RELATED WORK

Recent developments in deep learning have transformed image super resolution, first using CNN architectures and subsequently with GAN-based models for producing high- quality, realistic images. The methods have tackled problems such as noise removal, learning features, and image integrity, with implications for applications including surveillance and face recognition. Drawing on this research, our approach presents a customized GAN model with an innovative pre- training phase for the discriminator and denoising components in the generator, improving facial geometry and delivering improved inputs to methods such as GFPGAN (Wang et., al 2021). The next section discusses these advancements, contrasting their advantages and limitations, and illustrating how our technique improves super resolution research.

Goodfellow et al. (2014) presented a seminal adversarial setup for generative modeling, where two models a generative model representing the data distribution and a discriminator model separating real data from G's output are simultaneously trained in a minimax game. G is learned to maximize the probability that D makes an error, whereas D is learned to correctly identify real vs. generated samples. Trained with multilayer perceptrons, this model is entirely trainable by backpropagation without the need for sophisticated techniques such as Markov chains. Experiments by them illustrate that when the system is at equilibrium, G can restore the original data distribution and D learns to produce a constant output of 0.5, highlighting the effectiveness of their adversarial method. Ledig et al. (2017) proposed SRGAN, a single image super resolution framework utilizing a generative adversarial network to restore lost fine texture details during high upscaling. Contrary to error-minimizing pixel-wise approaches, SRGAN utilizes a perceptual loss consisting of adversarial and content losses to generate more photo-realistic images. Their method, which is based on a deep residual network, improves perceptual quality considerably, as verified by mean-opinion-score tests, rendering outputs closer to original high-resolution images. Lin and Shum (2004) examine the intrinsic limitations of restoring based super resolution algorithms (Z. Liu et., al.2022) (Wang et., al. 2018) that model the image formation process under local translation. They consider whether there are inherent limits to the resolution enhancement possible with these algorithms. Employing perturbation theory and a conditioning analysis of the coefficient matrix, the authors obtain explicit bounds on the super resolution Ledig et al. (2017) performance and find the minimum number of low resolution images required to achieve these limits. Their findings are reinforced by experiments on artificial and real data, which reveal insight into the real-world limits of reconstruction-based super resolution.

Savchenko et al. (2019) suggest a two-stage method for facial representation extraction to aid gender, identity and age recognition in images. The initial stage uses a customized MobileNet for face recognition, predicting age (as the average of top predictions) and gender as it extracts stable facial features. Hierarchical agglomerative clustering aggregates similar faces in the second stage, and collective predictions decide age and gender for each cluster. This technique, applied to an Android application, provides competitive clustering performance along with enhanced age/gender identification at reduced computational expenses.

Vijay et al. (2022) introduced a smart attendance system based on a channel-wise separable CNN to be used in group face recognition. The method derives image features and utilizes both SVM and Softmax classifiers for precise classification. Trained on the LFW database and tested with smart classroom data utilizing IoT-based edge computing, the system has an impressive 98.11% accuracy rate over other methods available in face recognition.

Zhang et al. 2018. present a Residual Dense Network for super resolution of images that maximizes hierarchical features within low resolution images. Contrary to most deep CNN architectures, which fail to tap into such features, RDN uses Residual Dense Blocks (RDBs) to harvest abundant local details via densely connected layers. RDBs utilize direct links among previous and the current layer to build an adjacent memory system, promoting fusion of features and maintaining training stability. The network then performs global feature fusion to combine these local features into a unified global representation. Findings from experiments conducted on standard datasets indicate that RDN performs better than state of the art methods under a variety of degradation conditions.

Abin Jose et al. 2018. introduce a content-based image retrieval system that employs convolutional layer features along with pyramid pooling to maintain spatial information and pro- duce compact, location-invariant feature vectors. Their system, experimented on Holidays and Oxford5K using AlexNet, performs better than fully connected layer or non-pooled convolutional feature-based methods.

3 GENERATIVE ADVERSIAL NETWORKS

Generative Adversarial Networks are a deep learning architecture proposed by Ian Goodfellow et al. in 2014. GANs fall under the category of generative models and are extensively applied to create high-quality artificial data such as images, videos, and text. GANs rely on a two-network structure involving a generator (G) and a discriminator (D), which engage in a zero-sum game paradigm. The generator attempts to produce data that is similar to actual data, whereas the discriminator attempts to separate real and generated data. This adversarial process propels both networks towards improvement, and highly realistic data generation results.

Mathematical Formulation: Generative Adversarial Networks are rooted in a two- network adversarial model with a generator G and a discriminator D. The generator seeks to produce authentic data samples, whereas the discriminator attempts to differentiate between genuine data. from generated samples. This is ex- pressed as a min-max optimization problem:

$$minmaxV(D,G) = Ex \sim pdata(x)[logD(x)] + Ez \sim pz(z)\left[log\left(1 - D\left(G(z)\right)\right)\right] \quad (1)$$

where:

- $x \sim pdata(x)$ implements actual data sampled from the actual data distribution.
- $z \sim pz(z)$ represents latent noise drawn from an earlier distribution (e.g Gaussian or Uniform).
- G(z) is the generator function that transforms noise into the data space.
- D(x) is the discriminator probability that x is real.

The discriminator aims to maximize the above objective, and the generator aims to minimize it, forming a two-player adversarial game.

4 METHODOLOGY

This section outlines our two-stage training procedure for image enhancement improvement with better facial recognition. The discriminator is pre-trained on the whole dataset in the first stage to learn the fine details and subtle features of real images. In the second stage, the complete GAN is trained, with an upgraded generator containing extra layers to preserve fine details and reduce noise. This approach guarantees that the generator generates high-quality images with important facial features, as confirmed by our experimental results.

4.1 Data Preparation

The data used in this research is a bespoke, non-public dataset, collected with strict regard to privacy and ethics since it comprises images of our colleagues taken under properly lit classroom settings to guarantee uniform lighting in all samples. The dataset contains 2,500 high-resolution images. of 19- to 24-year-old college students with a variety of facial expressions that cover the seven basic emotions: happy, sad, neutral, angry, frustrated, surprised, and a separate unique expression to provide full emotional coverage. Out of these 2,500 images, 2,000 pictures are utilized for training purposes and the remaining 500 images are left for testing. Each of the high resolution (HR) images, originally 128×128×3 in size, is subjected to a controlled degradation process to simulate low resolution (LR) inputs by first applying a Gaussian filter and then performing a down sampling operation to generate LR images of size 32×32×3. This intentional degradation not only creates a difficult input situation for the super resolution task but also preserves the important facial details and fine expression cues that are vital for subsequent facial expression recognition. The entire dataset is held in confidence to protect the privacy of the participating subjects, and all participants gave informed consent for their images to be used solely for research purposes within our institution.

4.2 Network Architecture

The network structure consists of three main elements: a generator, a discriminator, and a perceptual feature extractor. The generator is a deep convolutional network that aims to enlarge a low-resolution image (32×32) into a high resolution one (128×128) while preserving fine facial features. It starts with an initial convolutional layer with a large 9×9 kernel to extract low-level features, succeeded by a PReLU activation. Figure 1 show the Generator Architecture.

Figure 1: Generator architecture.

This is subsequently accompanied by a sequence of eight residual blocks improved, each block consisting of two dilated convolutional layers (dilation rate of 2) separated by batch normalization and PReLU activation. A channel attention mechanism is incorporated in every block to weigh feature channels dynamically, highlighting delicate facial expression signals. Skip connections within every residual block enable the combination of features at both low and high levels, thus allowing efficient gradient flow. Two sub-pixels up sampling blocks incrementally recover the spatial resolution at the end, and a final 9×9 convolutional layer reconstructs the output into a three-channel high resolution image.

The discriminator is designed to effectively differentiate between real high-resolution images and those produced by the network. It consists of eight convolutional layers with 64 filters initially and growing to 512 filters, and each of these uses 3×3 kernels and stride convolutions for down sampling the input in steps. LeakyReLU activations with an alpha value of 0.2 are used everywhere to ensure non-linearity and to prevent vanishing gradients. The convolutional stack is followed by two dense layers that summarize the learned features, leading to a sigmoid activation function that yields a probability score for the image's authenticity. This well-structured discriminator guarantees that the generator is constantly under pressure to make the super-resolved images even more realistic.

Accompanying the generator and discriminator is the perceptual feature extractor, which is constructed on top of a pre-trained ResNet101 model. By defining high-level se- mantic features from an intermediate layer (in this case," conv4 block23 out"), the extractor generates rich, detailed representations that assist in the perceptual loss. Overall training strategy exploits a number of loss functions, including adversarial (GAN) loss and discriminator loss to encourage realism, along with pixel-level and perceptual measures such as Mean Squared Error, Structural Similarity Index Measure, Peak Signal-to-Noise Ratio and (Hore´ and D. Ziou 2010). The PSNR and SSIM measures are especially significant since they quantitatively assess the fidelity and structural quality of the produced images. These losses collectively lead the generator not only to generate visually plausible high-resolution outputs but also to preserve the subtle details essential for follow-up facial expression recognition tasks. Figure 2 show the Discriminator Architecture.

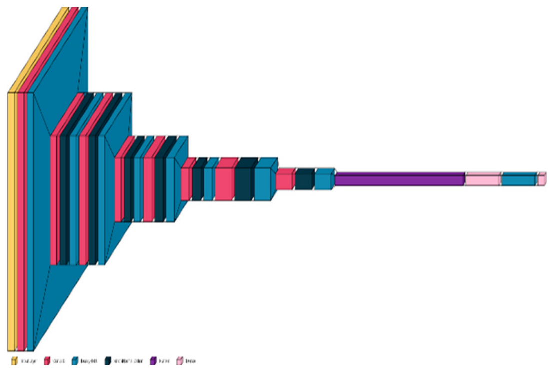

Figure 2: Discriminator architecture.

4.3 Loss Functions

To effectively train our super resolution network, we com- bine four loss functions together that jointly induce the model to produce high-quality, perceptually accurate images. Each loss term is focused on a particular feature of the quality of the image.

Discriminator Loss: This binary cross-entropy-based loss trains the discriminator to correctly identify between real high-resolution images and fake images produced by the model. It is expressed as:

$$LD = -EI_{HR} \_ \, pdata[logD(IHR)] - EI_{LR} \sim pdata\left[log\left(1 - D\left(G(ILR)\right)\right)\right] \quad (2)$$

Adversarial (GAN) Loss: This component of the loss encourages the generator to generate images that are indistinguishable from genuine images. It decreases the negative log the likelihood of the discriminator identifying the generated image as real:

$$LGAN = -EILR \sim pdata\left[logD\left(G(ILR)\right)\right] \quad (3)$$

Peak Signal to Noise Ratio Loss: Rather than using pixel-wise MSE directly, we use PSNR [18] as an image quality measure. where MAX represents the highest value of a pixel (usually 1 for normalized images or 255 otherwise). During our training, we would like to maximize PSNR; therefore, the PSNR loss is set to be the negative of PSNR:

$$\mathcal{L}_{\text{psnr}} = -10 \cdot \log_{10}\left(\frac{MAX^2}{\frac{1}{N}\sum_{i=1}^{N}|I_{HR}^{(i)} - G\left(I_{LR}^{(i)}\right)|^2}\right) \quad (4)$$

Structural Similarity Index Measure (SSIM) Loss: To maintain the perceptual quality and structural detail of the images, SSIM loss is used. It is a structure similarity measure between the real and the generated images:

$$LSSIM = 1 - SSIM\left(IHR, G(ILR)\right) \quad (5)$$

Training Strategy: The process of training is separated into two broad stages in order to guarantee stable convergence and strong learning of the fine details needed for both facial expression recognition and super resolution.

Phase 1: Discriminator Pre-training: First, the discriminator is trained on the whole dataset, which includes real HR images and their noisy or degraded LR counterparts. In this process, the discriminator learns the binary cross-entropy loss to differentiate between real HR images and initial generator outputs. This process provides the discriminator with a strong feature representation and sets a good foundation for the next adversarial training.

Phase 2: Adversarial Training: Following pre-training of the discriminator, the model proceeds to the adversarial training phase in which the generator and discriminator are alternately updated. This phase itself is split into two important sub-phases:

Generator updates: During this phase, the generator is learned such that it seeks to minimize an aggregate loss function, L_{SR}, as a weighted sum of several terms for loss. The Adversarial Loss (L_{GAN}) is responsible for training the generator to generate images that are capable of misleading the discriminator. The conventional Mean Squared Error loss is substituted with a Peak Signal-to-Noise Ratio Loss that prioritizes reconstruction fidelity by promoting greater PSNR values between the synthesized images and the ground-truth HR images. Additionally, the Structural Similarity Loss is used to maintain perceptual quality by emphasizing the structural details essential for correct facial expression recognition. This pairing guarantees that the generator outputs super-resolved images that are visually realistic as well as abundant in subtle details.

Discriminator update: Having set the parameters of the generator, the discriminator is then trained by minimizing its binary cross-entropy loss. This trains it to differentiate between actual HR images and super-resolved images produced by the network. By constantly improving its classification accuracy, the discriminator encourages the generator to generate increasingly realistic and detailed images. The trade-off updating of discriminator and generator creates a dynamic balance under which the generator is always nudged to generate better images and the discriminator is driven to better recognize subtle differences.

This balanced adversarial process, assisted by precisely calibrated optimizer parameters and learning rate schedules, is essential in obtaining high resolution outputs that preserve that fine features required for successful facial expressions recognition.

5 EXPERIMENTS AND RESULTS

The model is trained using an NVIDIA Quadro RTX 5000 GPU with 30GB VRAM and an Intel Xeon processor with 30GB RAM within an Anaconda environment via the Spyder IDE. Training was done across 100 epochs with a batch size of 128, utilizing TensorFlow libraries. Figure 3 show the Super

Resolution Output.

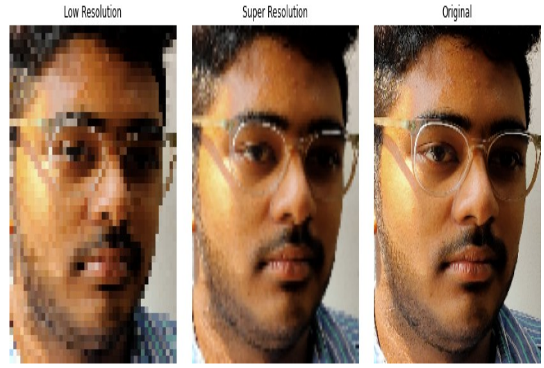

Figure 3: Super resolution output.

Figure 4: Super resolution output(noise).

Performance was assessed during training by using the discriminator loss (LD), generator loss (LG), and qualitative measures like Peak Signal-to-Noise Ratio and for measuring similarity in images Structural Similarity Index Measure Several experiments were conducted by modifying hyperparameters and applying different feature extraction methods to improve performance. Figure 4 show the Super Resolution Output(noise).

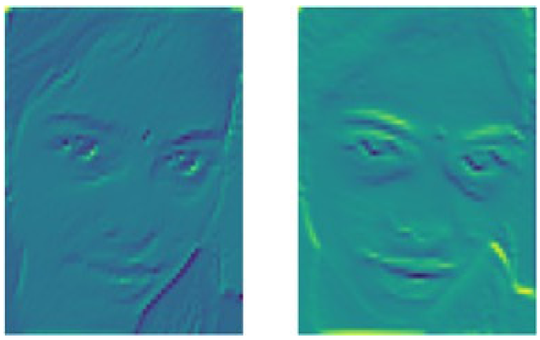

Figure 5: SRGAN Features.

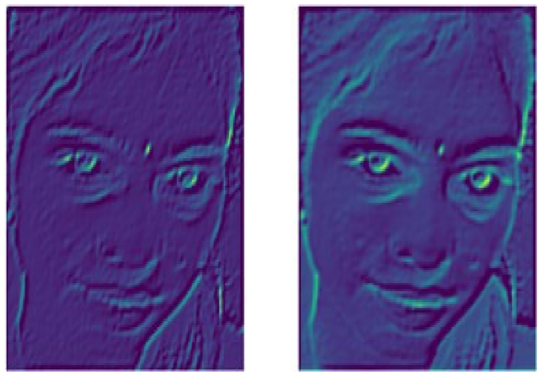

Figure 6: Proposed model features.

Figure 5 and 6 shows the SRGAN Features and Proposed Model Features respectively. First, the discriminator is separately pre-trained on another dataset prior to its inclusion in the GAN model.

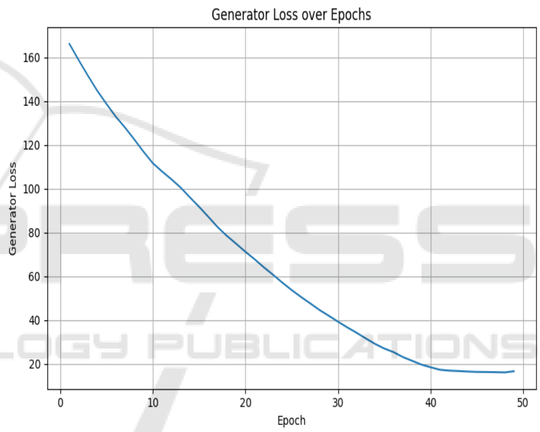

Figure 7: Generator training loss.

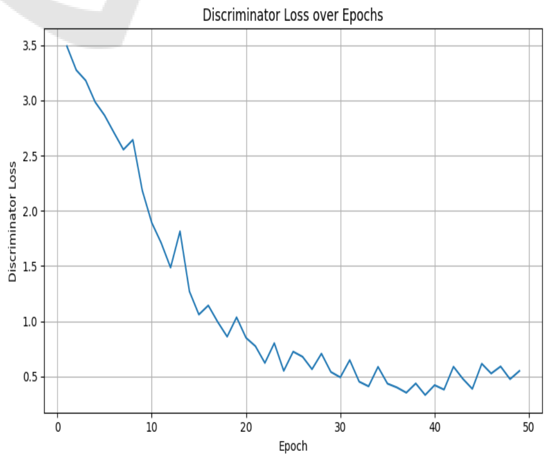

Figure 8: Discriminator training loss.

Figure 7 and 8 shows the Generator Training Loss and Discriminator Training Loss respectively.

A custom training loop is utilized, whereby in every epoch, the discriminator is trained both on fake images and on high resolution images in order to have a balance. At the same time, ResNet features are extracted and the GAN model is trained on batches with monitoring of all the metrics and losses. The training loop prints average statistics at the end of every epoch and saves the model every 10 epochs. Table 1 show the comparison of PSNR and SSIM across different methods.

Table 1: Comparison of Psnr and Ssim Across Different Methods.

Method	PSNR	SSIM
Bicubic	28.386	0.8249
EnhanceNet	28.548	0.8374
SFTGAN	29.913	0.8672
SRGAN	29.879	0.8694
Proposed Model	29.127	0.8264

Additionally, face images that are passed through the super resolution model are face reconstructed using the GFPGAN model to ensure recovered facial details. For reference, the following figure shows face reconstruction with and without super resolution.

Figure 9: Restored Image from Low Quality.

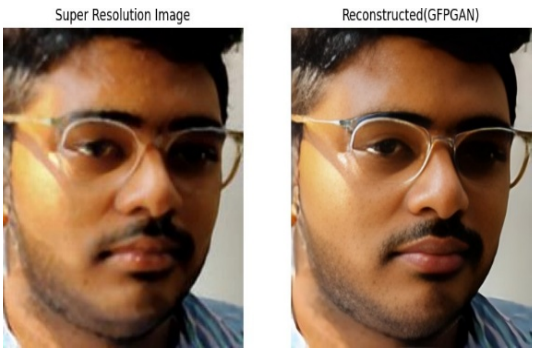

Figure 10: Restored Image from High Quality.

Figure 9 and 10 shows the Restored Image from Low Quality and Restored Image from high Quality respectively.

6 CONLUSIONS AND FUTURE SCOPE

In this work, we proposed a custom generative adversarial network to increase image resolution without compromising fine facial details essential for expression recognition. Our method combines a deep generator with enhanced residual blocks and channel attention, a stable discriminator, and a ResNet101-based perceptual feature extractor. With the use of a combination of adversarial, PSNR, and SSIM loss functions, the model efficiently produces high-quality super-resolved images while retaining vital facial details. Experimental verification on a varied dataset, recorded under standardized lighting with high variability in facial expressions, validates the efficacy of our approach for reconstructing high-fidelity images for facial analysis applications.

Future work will involve improving attention mechanisms to better capture feature extraction, generalizing the model to video super resolution, and enhancing generalization across multiple datasets. It will also investigate optimizing the model for real-time processing on edge and mobile devices to widen its applications in practice. These enhancements are expected to further boost the efficiency and adaptability of the developed model for real-world applications in facial recognition and other related tasks.

REFERENCES

A. Hore´ and D. Ziou,” Image quality metrics: PSNR vs. SSIM,” in Proceedings of the 2010 IEEE International Conference on Pattern Recognition (ICPR), pp. 2366-2369, 2010, doi: 10.1109/ICPR.2010.579.

A. Jose, R. D. Lopez, I. Heisterklaus and M. Wien,” Pyramid Pooling of Convolutional Feature Maps for Image Retrieval,” 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 480- 484, doi: 10.1109/ICIP.2018. 8451361.

A. A. Abello and R. Hirata,” Optimizing Super Resolution for Face Recognition,” 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, 2019, pp. 194-201, doi: 10.1109/SIBGRAPI.2019.00034.

Alvarez-Ramos, V. & Ponomaryov, V. & Sadovnychiy, Sergiy. (2018). ” Image super resolution via Wavelet

Feature Extraction and Sparse Rep- resentation,” Radioengineering, 27, 602- 609, 10.13164/re.2018.06 02.

C. Ledig et al.,” Photo-Realistic Single Image super resolution Using a Generative Adversarial Network,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 105-114, doi: 10.1109/CVPR.2017.19.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Genera- tive Adversarial Networks, 2014, https://arxiv.org/abs/1406.2661.

L. Hui and S. Yu-jie,” Research on face recognition algorithm based on improved convolution neural network,” 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, 2018, pp. 2802- 2805, doi: 10.1109/ICIEA.2018.8398 186.

M. Elad and M. Aharon,” Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries,” IEEE Transactions on Image Processing, vol. 15, no. 12, pp. 3736-3745, Dec. 2006, doi: 10.1109/TIP.2006.881969.

Pabba, C., & Kumar, P. (2022).” An intelligent system for monitor- ing students’ engagement in large classroom teaching through facial expression recognition,” Expert Systems, 39(1), e12839, https://doi.org/ 10.1111/exsy. 12839.

Pabba, Chakradhar & Kumar, Praveen. (2023).” A vision-based multi- cues approach for individual students’ and overall class engagement monitoring in smart classroom environments,” Multimedia Tools and Applications, 83. https://doi.org/10.1007/s11042-023-17533-w.

Savchenko AV.” Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet,” PeerJ Computer Science 5:e197, 2019, https://doi.org/10. 7717/peerj-cs.197.

Shi, Wenzhe & Caballero, Jose & Husza´r, Ferenc & Totz, Johannes & Aitken, Andrew & Bishop, Rob & Rueckert, Daniel & Wang, Zehan. (2016).” Real-Time Single Image and Video super resolu- tion Using an Efficient Sub-Pixel Convolutional Neural Network.” https://doi.org/10.48550/arXiv.1609.05158.

V. M., D. R. and P. B. S.,” Group Face Recognition Smart Attendance System Using Convolution Neural Network,” 2022 International Confer- ence on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2022, pp. 89-93, doi: 10.1109/WiSP- NET54241.2022.9767128.

Wang, Xintao & Yu, Ke & Wu, Shixiang & Gu, Jinjin & Liu, Yihao & Dong, Chao & Loy, Chen Change & Qiao, Yu & Tang, Xiaoou. (2018). ESRGAN: Enhanced super resolution Generative Adversarial Networks. https://doi.org/10.48550/arXiv.1809.00219.

Wang, Xintao, et al.” Towards real-world blind face restoration with generative facial prior.” Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

Xiao H, Wang X, Wang J, Cai JY, Deng JH, Yan JK, Tang YD.” Single image super resolution with denoising diffusion GANS,” Sci Rep. 2024 Feb 21;14(1):4272. doi: 10.1038/s41598-024-52370-3. PMID:38383573; PMCID: PMC11222509.

Z. Liu, Z. Li, X. Wu, Z. Liu and W. Chen,” DSRGAN: Detail Prior-Assisted Perceptual Single Image super resolution via Generative Adversarial Networks,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 11, pp. 7418-7431, Nov. 2022, doi: 10.1109/TCSVT.2022.3188433.

Zhang, Yulun and Tian, Yapeng and Kong, Yu and Zhong, Bineng and Fu, Yun,” Residual Dense Network for Image super resolution,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

Zhouchen Lin and Heung-Yeung Shum,” Fundamental limits of reconstruction-based superresolution algorithms under local translation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 83- 97, Jan. 2004, doi: 10.1109/TPAMI.200 4.1261081.