IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PIXEL PERFECTION: SRGAN BASED IMAGE ENHANCEMENT

[1] Prasath B, [2] Ramakrishna KS, [3] Pradeep Raj M, [4] Karthickeyan LK

[1] Assistant Professor (Sr.G.)

[1] Department of Artificial Intelligence and Data Science,

[1] KPR Institute of Engineering and Technology,

Coimbatore, India.

Abstract: In order to improve picture resolution for a variety of applications, this research examines how well Super-Resolution Generative Adversarial Networks (SRGANs) operate. To increase visual quality and detail preservation in the produced high-resolution (HR) pictures, we present a modified SRGAN architecture that incorporates perceptual loss and residual blocks. By comparing the suggested method's performance to a benchmark picture dataset, quantitative measurements such as Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) show that it can outperform conventional interpolation techniques. These outcomes show how SRGANs may be used to further picture enhancement and restoration in a variety of fields. We also discuss ways this promising technology may go in the future and possible enhancements.

Keywords **–** Generative adversarial networks, Blurred image, Restoration, Enhancing, GAN, SRGAN.

I. Introduction

Images that are blurry due to motion, out-of-focus, or noise greatly reduce their usefulness in security, remote sensing, and medical imaging, among other applications. Conventional restoration techniques are generally insufficient for handling various blur kinds, even when they work well for some problems like denoising or dehazing and quality, thereby improving the effectiveness of these systems.

In this study, a new method for picture enhancement and restoration using Super-Resolution Generative Adversarial Networks (SRGANs) is proposed. Through the optimization of the SRGAN architecture, we hope to overcome the shortcomings of current techniques to handle blurry pictures.

The main objective of this optimization is to provide realistic and aesthetically pleasing images with restored details that perform better than the state-of-the-art methods in terms of visual quality and quantitative evaluations. Our refined SRGAN method shows encouraging outcomes for improving and restoring blurry photos, opening the door for developments in deep learning-based image restoration.

II. Literature Survey

Image restoration and enhancement techniques have garnered significant attention across various domains, ranging from medical imaging to remote sensing and surveillance systems.

This literature survey provides an overview of recent advancements in this field, focusing on key research works utilizing deep learning models, particularly Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs)Yuan (2023) proposes an optimization technique for blurry image repair and enhancement on Super-Resolution Generative Adversarial Networks (SRGAN), exploring its effectiveness in improving image quality and fidelity through adversarial training [1]. Building upon this, Xu et al. (2022) introduce an image enhancement algorithm leveraging GAN neural networks, highlighting the potential of GANs in enhancing image features and textures for improved visual quality [2]. This progression underscores the continuous evolution and exploration of GAN-based approaches in image restoration and enhancement tasks. Furthermore, Yamashita and Markov (2020) focus on enhancing medical images with super-resolution techniques, emphasizing the importance of high-resolution images for accurate diagnosis and analysis in healthcare applications [3].

Dai et al. (2018) complement these findings by demonstrating the effectiveness of Convolutional Neural Networks, particularly SRCNN, in improving the quality of radar images for better detection and analysis in radar-based applications [4].Finally, Sophia et al. (2022) propose an efficient method for blind image restoration using GANs, showcasing the potential of GAN-based approaches for addressing real-world image restoration challenges without prior knowledge of the degradation process [5].

SRGAN combines a generator and discriminator network, transforming LR images to HR outputs and evaluating image quality for feedback. It excels in image recognition, deblurring, and denoising, showing promise in restoration tasks. Deep learning-based methods are pivotal for improving degraded images, with SRGAN leading the charge in enhancing perception and fidelity.

III. Methodology

The suggested method uses a framework for image enhancement that is based on GANs. An image with poor quality or incompleteness is processed by the generator network, which produces an image with improved quality and more details and fidelity. The discriminator network feeds back information on image quality to the generator network, which helps the generator network refine its output. The framework consists of two main components: a generator (G) that generates synthetic samples in order to fool the discriminator and a discriminator (D) that determines which samples are created and which are legitimate. In the training phase, the generator creates counterfeit images that mimic the original input, and the discriminator uses a gaming approach to identify the counterfeit images.

SRGAN

SRGAN represents a super-resolution network proficient in producing adversarial networks. The architecture of SRGAN network is illustrated in Figure 1.

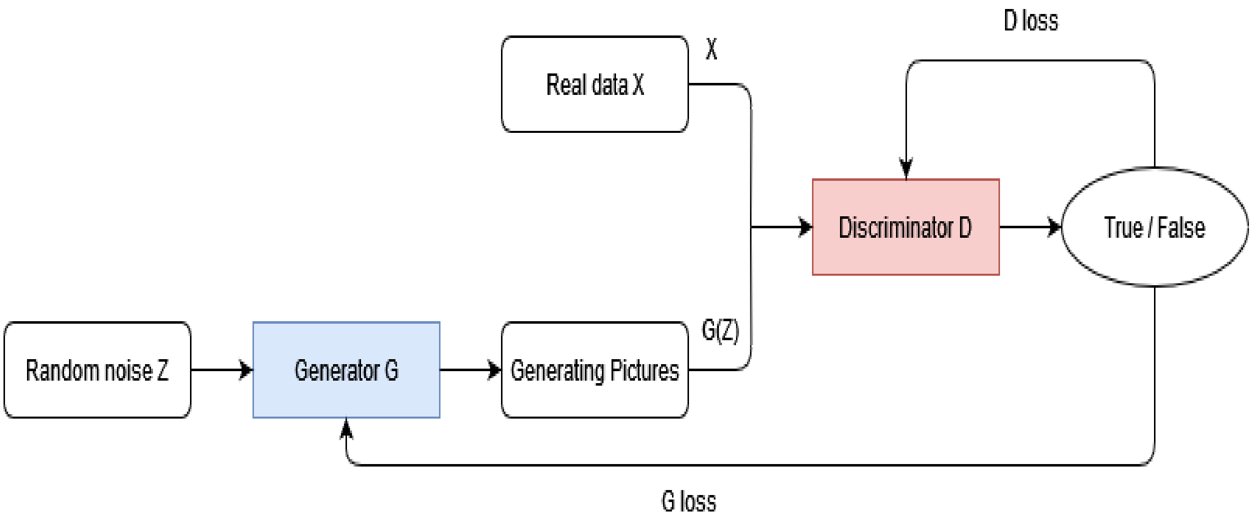


Figure 1. Structure of the SRGAN network

In SRGAN network structure, random noise Z serves as input to the Generator G, producing the generated image $G(Z)$. Mathematically, this procedure can be shown as

$$G: Z \rightarrow G(Z) \quad (1)$$

Subsequently, both the real image X and the generated image $G(Z)$ are fed into the Discriminator D, which outputs a boolean value indicating the authenticity of the images. Mathematically, the Discriminator can be represented as

$$D: \{X, G(Z)\} \rightarrow \{True, False\} \quad (2)$$

The Discriminator incurs a loss, denoted as L_D, which is typically defined using binary cross-entropy or other appropriate loss functions. Similarly, the Generator incurs a loss, denoted as L_G, which promotes the Discriminator to classify the created images as real. A min-max game is used to mathematically design this adversarial training framework.

$$\underset{G}{\text{Min}}\,\underset{D}{\text{Max}}\, V(D,G) = \mathbb{E}_{X\sim P_{data}(x)}[\log D(X)] + \mathbb{E}_{z\sim P_z(z)}[\log(1 - D(G)))] \quad (3)$$

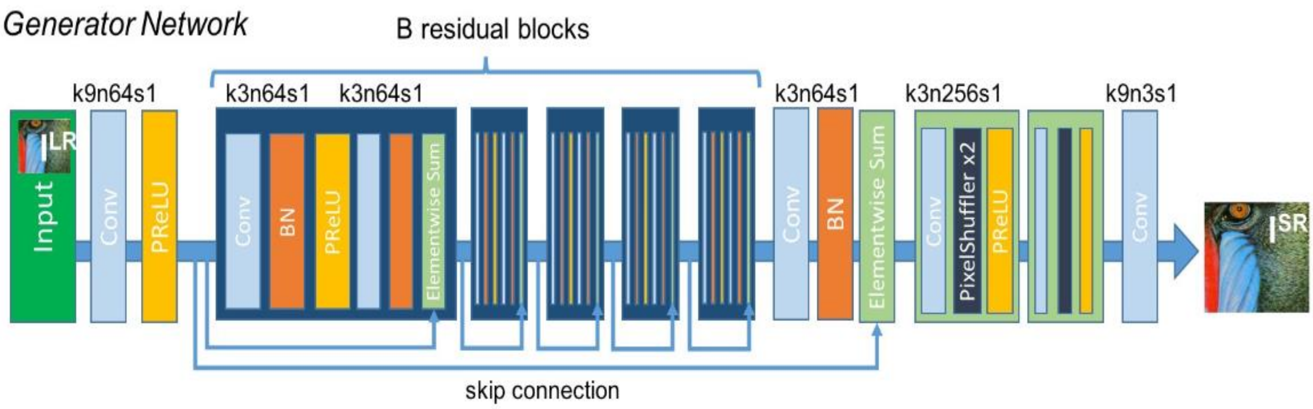
Where $p_{data}(X)$ represents the distribution of real images and $p_z(Z)$ represents the distribution of the input noise Z. This formulation encourages the Discriminator to correctly classify real and generated images, while simultaneously encouraging the Generator to produce images that are classified as real by the Discriminator.

Network Architecture

The adversarial min-max problem can be expressed as follows through the alternate training of a generator network and a discriminator network. The primary objective of the generator network is to generate high-quality images that are hard for the discriminator network to distinguish from actual HR pictures. On the other hand, the discriminator network seeks to accurately distinguish between real high-resolution photos and produced images. The first term in this equation is the discriminator network's capacity to forecast the log likelihood of properly categorizing real HR pictures as authentic.

Generator Network

A very deep generator network G, represented by B residual blocks with a regular architecture, is located in the center of Figure 2. Each block consists of two 64 feature map convolutional layers with smaller 3 × 3 kernels along with batch-normalization layers before being activated by ParametricReLU. To improve the



resolution of the input image, two trained sub-pixel layers of convolution are subsequently added.

Figure 2. Generator network design

Discriminator Network

SRGAN's discriminator network is trained to distinguish between real high-resolution (HR) images and artificial super-resolution (SR) samples. The architecture is carefully designed to prevent max-pooling across the network, as shown in Figure 3. Alternatively, an activation function of LeakyReLU with a slope of 0.2 (α=0.2). The network is trained to handle the maximizing problem. It consists of eight convolutional layers, each of which has 3 filter kernels. Each layer has 512 kernels total, a progressive increase from 64.

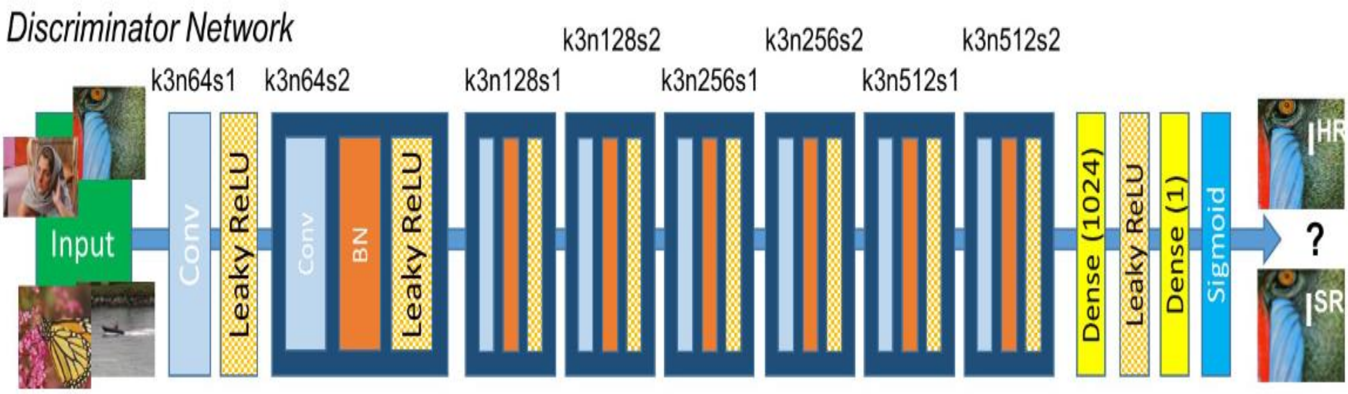


Figure 3. Discriminator network design

IV. Training and Testing

We train our proposed model on the VOC2012 dataset. The dataset's low-resolution (LR) and high-resolution (HR) images are scaled by a factor of 4 in each experiment. We evaluate the performance of the SRGAN method in comparison to the SRCNN approach using two distinct test datasets, set5 and set14. During the testing phase, we primarily focus on three key indicators:

- Structural Similarity Index Measure(SSIM)
- Visual Effect
- Peak Signal-to-Noise Ratio(PSNR)

Specifications and parameters for training

This project leverages an NVIDIA GeForce RTX 3090 Ti graphics card for hardware acceleration and utilizes Python for development. The development environment is set up within Visual Studio Code (VSCode), while PyTorch serves as the deep learning framework for experimentation. The project downsamples high-resolution (HR) images (represented in BGR color format with 3 channels) into low-resolution (LR) images using a bicubic kernel with a downsample factor of 4. During training, batches consisting of 16 randomly selected 96x96 HR sub-images are extracted from various training images.It should be mentioned that any size photo may be utilized with the generator model because it is fully convolved. The range of the HR photos is [-1,1], but the scaled range of the LR input images is [0,1]. Therefore, the L1 loss is computed using the image intensity range [-1,1]. Adam was used for optimization, with β1=0.9. In this experiment, the learning rate is set at 10-4, the discriminator has 8 convolution layers, and the generator has 16 residual blocks. Throughout the testing, batch normalization updates were disabled in order to obtain a result that solely depended on the input deterministically.

V. Analyzing Results

Analysing Training and Validation Dataset Results

The training set's batch size is 64, the upscale parameter is 4, and the experiment consists of 120 epochs. The PSNR and SSIM values of the validation set are obtained by computing the mean square error between the HR and SR shots after the optimizer has processed the input pictures. Following 120 epochs, the validation set's PSNR and SSIM values were determined. It is easy to observe that the PSNR and SSIM of the validation set are both increasing gradually and steadily as the number of epochs rises, even if there were a few little discrepancies at first. The first epoch of these has a PSNR of around 21 dB, and after 120 epochs, it increases to about 23.6 dB, at which point the human eye can immediately discern between SR and HR images.

On the other hand, the validation set's SSIM grew during the training, varying from 0.64 at the start to around 0.78 at the finish. It is evident that the produced SR pictures are becoming crisper and of greater quality as a result of ongoing training and optimization. Furthermore, it is anticipated that after additional training epochs, the SSIM and PSNR values for the validation set would rise even further.

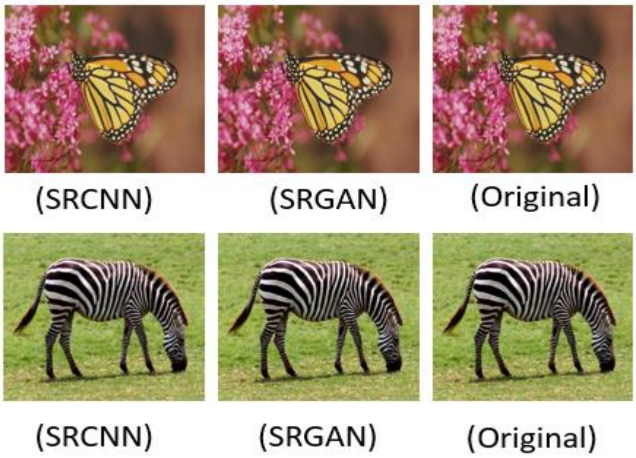
Comparison between SRGAN and SRCNN

Performance comparison between SRCNN and SRGAN is conducted using HR images as references, with the datasets Set5 and Set14. This analysis aims to delineate the disparities between the two methods across these datasets.

Table 1. Comparison of PSNR and SSIM of SRGAN and SRCNN on Set5 and Set14 datasets

Set5	SRGAN	SRCNN
PSNR	30.67	30.22
SSIM	0.894	0.873
Set14	**SRGAN**	**SRCNN**
PSNR	27.23	26.98
SSIM	0.831	0.794

Table 1 shows that SRGAN performs better than SRCNN in terms of SSIM and PSNR values. In comparison to SRCNN, SRGAN achieves greater PSNR and SSIM values, suggesting superior picture quality. This advantage is ascribed to SRGAN's use of a perceptual loss function, which motivates the generator network to generate pictures with better resolution and greater resemblance to actual images, leading to outputs that are crisper and more lifelike. To further improve its capacity to retain small details, SRGAN also uses an element loss function that compares features retrieved from the produced and ground truth pictures. Additionally, SRGAN incorporates an adversarial loss function, which further encourages the generation of realistic, artifact-free images. These combined factors contribute to SRGAN's ability to produce high-quality



images with fewer artifacts compared to SRCNN. Figure 4 provides a detailed visualization of the specific improvements in training outcomes.

Figure 4. Comparison of SRGAN and SRCNN on the same image.

VI. Conclusion

In summary, this study presents a state-of-the-art method based on the SRGAN architecture for reconstructing and enhancing fuzzy photos. The recommended technique effectively enhances the features and textures of blurry photos while maintaining their natural appearance. The technique combines deep convolutional neural networks with adversarial and total variation loss functions to produce visually realistic and aesthetically pleasing images..

In experimental evaluations on several kinds of hazy photographs, the suggested approach performs better than current technologies, exhibiting improved quantitative measurements and visual quality. The method generates more aesthetically pleasing images than SRCNN, with higher PSNR and SSIM ratings. The suggested technique may find use in remote sensing, medical imaging, security imaging, and other fields. It may also be expanded to tackle issues with super-resolution, deblurring, and denoising in picture restoration activities.

In conclusion, the suggested approach offers a workable way to maximize the restoration and improvement of blurry photos, marking a substantial breakthrough in the field of image restoration with deep learning approaches. This study might stimulate greater investigation in this area and lead to the creation of more potent picture restoration techniques..

References

[1] Yuan, Ziqi. "Restoration and enhancement optimization of blurred images based on SRGAN." In Journal of Physics: Conference Series, vol. 2664, no. 1, p. 012001. IOP Publishing, 2023.

[2] Xu, Bo, Dong Zhou, and Weijing Li. "Image enhancement algorithm based on GAN neural network." IEEE Access 10 (2022): 36766-36777.

[3] Yamashita, Koki, and Konstantin Markov. "Medical image enhancement using super resolution methods." In Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part V 20, pp. 496-508. Springer International Publishing, 2020.

[4] Dai, Yongpeng, Tian Jin, Yongping Song, Hao Du, and Dizhi Zhao. "SRCNN-based enhanced imaging for low frequency radar." In 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama), pp. 366-370. IEEE, 2018.

[5] Sophia, SG Gino, Syed Umar, and N. Visvas. "An efficient method for Blind Image Restoration using GAN." In 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), pp. 1-8. IEEE, 2022.

[6] Liu, Ming-Yu, Thomas Breuel, and Jan Kautz. "Unsupervised image-to-image translation networks." Advances in neural information processing systems 30 (2017).

[7] Zhang, Kai, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising." IEEE transactions on image processing 26, no. 7 (2017): 3142-3155.

[8] Dong, Chao, Chen Change Loy, Kaiming He, and Xiaoou Tang. "Learning a deep convolutional network for image super-resolution." In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13, pp. 184-199. Springer International Publishing, 2014.

[9] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee. "Accurate image super-resolution using very deep convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1646-1654. 2016.

[10] Lim, Bee, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. "Enhanced deep residual networks for single image super-resolution." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 136-144. 2017.