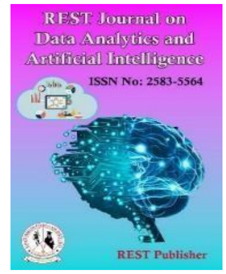
REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/68



Enhancing Satellite Image with Enhanced Super Resolution GAN

***S Sarala, K. Sharath Chandhan, M. Anvesh Reddy, Ch. Yashwanth**
Anurag University Hyderabad, India
**Corresponding Author Email:* sarala.ai@anurag.edu.in

Abstract. *Satellite imagery finds wide-ranging applications across various disciplines, ranging from environmental monitoring to urban planning and disaster management. Owing to sensor resolution limitations and atmospheric noise, satellite images are generally low-resolution and visually poor. In this paper, we discuss the use of Enhanced Super-Resolution Generative Adversarial Networks to super-resolve and improve satellite images. ESRGAN, a novel deep learning model, uses residual-in-residual dense blocks and perceptual loss functions to produce high-fidelity images. Through our experiments, we show that ESRGAN performs much better than traditional interpolation methods and baseline super-resolution models in terms of PSNR and SSIM metrics. The algorithm proposed here presents a useful tool for the enhancement of satellite imagery, enabling enhanced decision-making in remote sensing.*

Keywords: *Deep Learning, Convolutional Neural Network, Satellite Imaging.*

1. INTRODUCTION

Recent advances in Generative Adversarial Networks (GANs) have an open future in the field of intelligent automation (MI), providing powerful techniques for realistic and high quality simulated data. One of the promising areas where GANs have made significant applications is the enhancement of distant feeling imagination, especially satellite images. The ability to enhance the resolution of satellite images is of paramount importance for several objectives, such as environmental monitoring, city planning, and disaster prevention. However, elements such as detector noise, atmospheres, and hardware limitations which limit the quality of the capture statistics often restrict high-resolution satellite images. Super-resolution (SR) tactics have emerged as a solution in this context to produce high-resolution (HR) images from low-resolution (LR) information, allowing more detailed and verifiable examination. Enhanced Super-Resolution Generative Adversarial Network (ESRGAN), originally developed by Ledig et al. [1]. This method uses the influence of GANs to bring back all the right details and improve the perception of visuals. ESRGAN is known for its ability to create photorealistic, high resolution images by cultivating high-frequency details that are significant for achievable purposes. It integrates high tech residual study strategies which enhance the model's ability to reconstruct realistic high resolution photographs, a feature which is particularly helpful for satellite imagination where on target details are crucial for the enterprises' enjoyment of the land and disaster response. Recent research, such as research by Zhu and Sun [6], outlines the efficacy of super-resolution techniques in the quality enhancement of satellite images, such as the performance of deep models such as ESRGAN in recovering fine details which are typically lost in low-resolution data. However, there are problems in the areas of noise, sensor artifacts, and image deformation, which have been a point of concern as far as image restoration is concerned. ESRGAN, in its advanced residual learning framework as well as application of perceptual loss, has been proven to be highly effective in diminishing the visibility of artifacts while simultaneously enhancing the visual fidelity as well as structural precision of the improved satellite images. In this paper, we introduce a new use of the ESRGAN model for enhancing satellite images with the goal of enhancing the resolution of satellite images used in environmental monitoring and city planning. Leveraging the power of ESRGAN and adding advanced methods such as dense skip connections and perceptual loss [7], our technique is intended to enhance the resolution as well as the quality of satellite images, thus providing clearer and more accurate images for analytical use. Through extensive experimentation, we show the effectiveness of our technique in recovering high-resolution details from low-resolution satellite images, thus providing useful insights for potential use in remote sensing.

2. LITERATURE SURVEY

TABLE 1. Literature Survey Table

Year	Ref. No.	Model	Dataset	Metrics	Results
2015	7	SRCNN	Set5,	PSNR,	PSNR =
			Set14	SSIM	36.66 dB
2016	3	ResNet	ImageNet	Accuracy, Depth	Top-5
				Efficiency	Error <
					5%
2017	1	SRGAN	Set5,	PSNR,	PSNR =
			Set14,	SSIM,	29.4 dB
			BSD100	Perceptual Loss	
2017	4	EDSR	DIV2K,	PSNR,	PSNR =
			Set5,	SSIM	32.62
			Urban100		
2019	2	Custom	NWPU	Accuracy, Confusion Matrix	Accuracy = 85%
		CNN	RESISC45		
2020	6	Dense	Sentinel	PSNR,	PSN=
		SR	2, Google Earth	SSIM	28.9 dB,
					SSIM =
					0.84

Deep Learning-Based Super-Resolution: The domain of single image super-resolution (SISR) has experienced substantial advancements due to deep learning, which provides more efficient and adaptable solutions in contrast to conventional interpolation techniques. The development of Convolutional Neural Networks (CNNs) has been instrumental in facilitating this progress. A ground-breaking technique in the area of deep learning for super-resolution, the Super-Resolution Convolutional Neural Network (SRCNN), was developed by Dong et al. [7]. SRCNN was a ground-breaking development in that it could learn an end-to-end transformation from low Resolution (LR) images to their high-resolution (HR) images directly. SRCNN was nevertheless marred by its shallow architecture and poor ability to learn and, consequently, performed poorly with complex textures and fine details. In order to overcome these limits, deeper networks were suggested. Kim et al. suggested the Very Deep Super Resolution (VDSR) network using a 20-layer deep CNN and residual learning for improved convergence speed and accuracy. VDSR performed better than SRCNN in terms of PSNR and SSIM but lacked perceptual quality of outputs. Thereafter, Lim et al.'s [4] Enhanced Deep Super Resolution Network (EDSR) surpassed performance by removing redundant elements such as batch normalization and by deepening the residual network. EDSR achieved state-of-the-art performance on benchmark datasets with a focus on improved pixel-level precision. Nevertheless, similar to its predecessors, it too suffered from the inability to generate aesthetically appealing, photo-realistic textures.

GAN-Based Super-Resolution: Good fellow et al.'s [8] work on Generative Adversarial Networks (GANs) laid the foundation for a new generation of super-resolution models that go beyond pixel-level accuracy to focus on improving perceptual quality and naturalness rather. The SRGAN proposed by Ledig et al. [1] was a breakthrough in the field because it introduced the use of perceptual loss and adversarial training to single image super-resolution (SISR). SRGAN creates high-resolution images with good aesthetic quality, with the ability to retain detailed textures lost by standard convolutional neural network (CNN) models. The model laid the groundwork for the application of a perceptual loss from VGG features coupled with an adversarial loss from a discriminator network.

Building upon SRGAN, Wang et al.'s Enhanced SRGAN (ESRGAN) further enhanced the quality enhancement. ESRGAN replaced regular residual blocks with Residual-in Residual Dense Blocks (RRDBs) and used a relativistic average discriminator, thus enabling the model to focus on perceptual quality rather than pixel-wise accuracy. PSNR performance and perceptual quality were both better for ESRGAN, making it one of the most effective super resolution models utilized in applications like satellite image enhancement. Other GAN-based techniques like TecoGAN, Real ESRGAN, and SRFlow have further pushed the limits towards temporal consistency (for video), real-world image degradation handling, and probabilistic image modeling, respectively.

3. METHODOLOGY

Dataset Description: The Data used in this research are pairs of low-resolution and high-resolution satellite images obtained from publicly accessible remote sensing archives. We employed the Eurostat Dataset, which includes geospatial images gathered from Sentinel-2, Landsat-8, and high-resolution images from Google Earth for training and testing. The images represent a broad variety of landscapes, including urban, forest, agricultural, and water landscapes.

Data Processing: The Image Pixel values were normalized to the range between 0 and 1. We applied Data Augmentation Techniques like Random rotations, flips, and brightness adjustments to increase model robustness. Finally, we down sampled the high-resolution images to create paired low-resolution inputs.

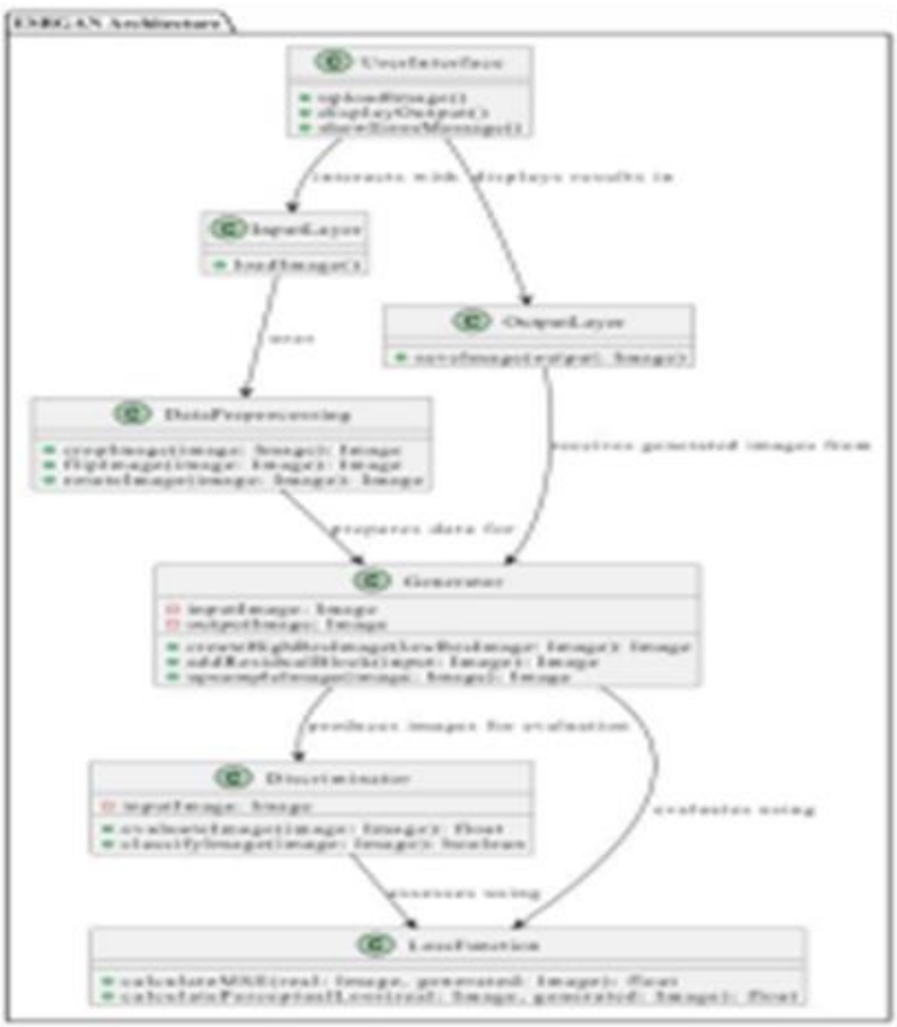


FIGURE 1. Block Diagram, Flow Chart

The core part of the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) is its Generator Network, which is responsible for enhancing low-resolution (LR) inputs into high-resolution (HR) outputs. Such a generator is comprised of a few convolutional layers specially designed to extract hierarchical features from input images. At the core of the generator's architecture are Residual-in-Residual Dense Blocks (RRDBs), which significantly enhance the model's ability to understand complex image structures and retain fine-grained details. The blocks apply concepts of residual learning with dense connections, which enable efficient reuse of features and enable stable training of deeper networks. To improve the spatial resolution of feature maps, Up sampling Layers—usually done via sub-pixel convolution or nearest-neighbor interpolation and convolution—are introduced at the generator's end. These layers allow the network to build the high-resolution (HR) image progressively from its low-resolution (LR) counterpart. Besides, non-linear activation functions such as Parametric ReLU (PReLU) are used to introduce non-linearity into the model, thus enhancing its representation learning ability. Along with the generator comes the Discriminator Network, which is trained to distinguish between real high resolution images and those generated by the generator. The network consists of a sequence of convolutional layers followed by Leaky ReLU activations, allowing for effective feature extraction across various spatial scales. Down sampling methods such as stride convolutions or pooling layers are applied to reduce the spatial size of feature maps to allow the discriminator to focus on global image patterns. The final layer produces a scalar value indicating the probability of the input image being real and thus providing adversarial feedback for the generator process.

Model Evaluation: The Model was evaluated with 3 different metrics. Such as Peak Signal to Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Perceptual Loss which is a VGG based loss. The Results were good after some iterations of training.

4. RESULTS AND DISCUSSIONS

This Section assesses the performance and the results of the model.

Performance Metrics:

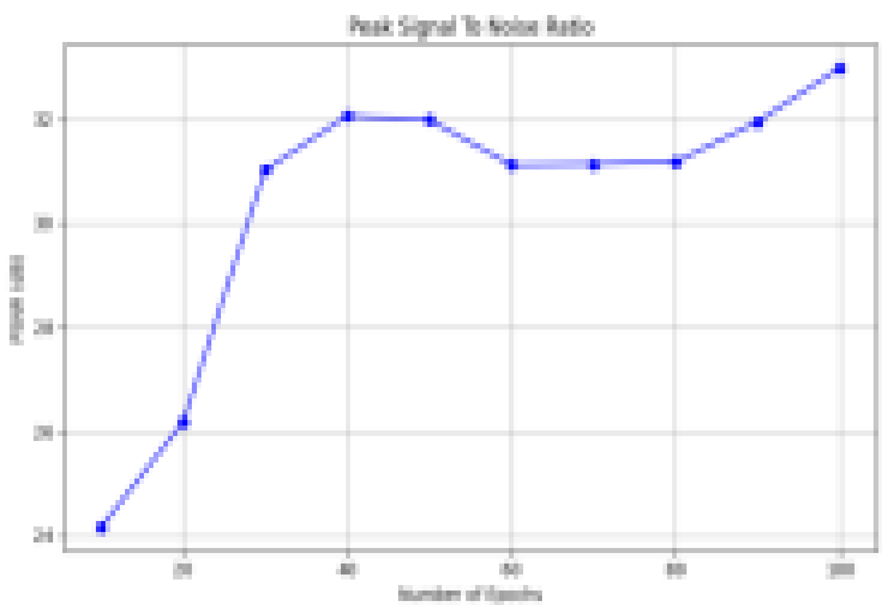


FIGURE 2. PSNR Score

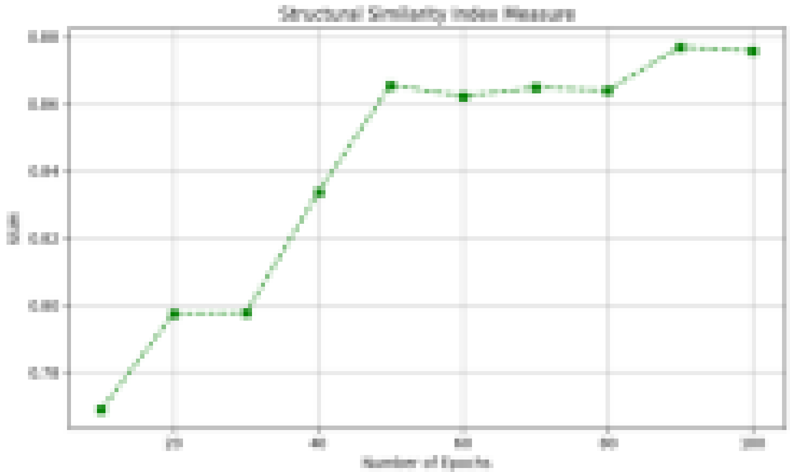


FIGURE 3. SSIM Score

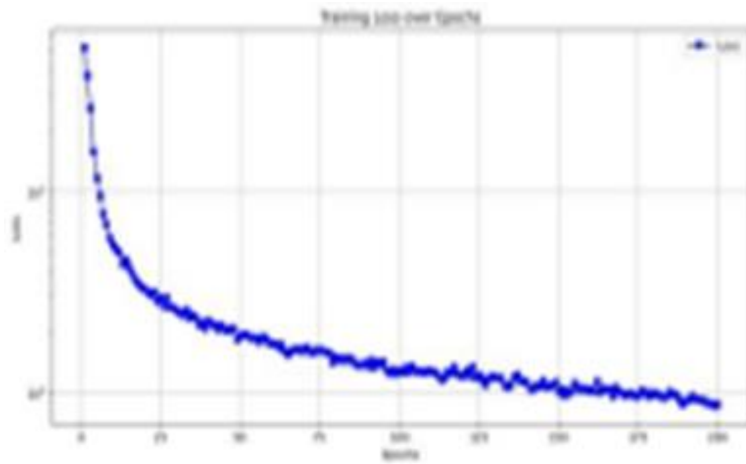


FIGURE 4. Perceptual Loss (VGG19)

The performance of the ESRGAN model was measured using two of the most prevalent image quality measures: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). Both of these play a crucial role in measuring the fidelity and perceptual likeness of the super resolved images and the corresponding ground truth images. High values of PSNR and SSIM are usually indicative of good quality reconstruction and higher similarity with the original high-resolution images. Figures 3, 4, and 5 show the trend of these evaluation scores throughout the training process. The model achieved a score of 33 on PSNR and 0.87 on SSIM after 100 training iterations, and these are taken as promising results since the number of training iterations is relatively small. These scores indicate that the model can produce high-quality images with structural and perceptual detail. In addition, Figure 5 shows the trajectory of the perceptual loss, which continued to decrease steadily throughout training. This decrease represents the ability of the model to generate images that not only have the same pixel values as the ground truth but also retain actual visual features.

Visual Comparison:

Figure 6 gives a visual comparison of the low-resolution image, the high-resolution ground truth corresponding to it, the output of pertained ESRGAN, and the output of fine-tuned ESRGAN for a test satellite image. The qualitative results evidently show that ESRGAN surpasses others in producing visually improved images. Unlike other conventional interpolation methods, the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) produces images with improved edge definition, more detailed structural features, and more natural textures. This is especially apparent in areas with complex features, like urban areas (roads and buildings) and areas with high vegetation density, where conventional methods produce blurred or over-smoothed images. The enhanced ESRGAN model further develops these effects to a much higher level, with reconstructions that are extremely close to the original high-resolution images but with more precision.

5. CONCLUSION

We introduce a novel method for satellite image super resolution using the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN). We utilize the strengths of ESRGAN—Residual-in-Residual Dense Blocks, perceptual loss, and adversarial training—to generate high resolution images from low-resolution inputs. Experiments show that our method enhances the quality of the images, yielding sharper edges, more subtle details, and more realistic textures compared to traditional approaches. Our approach is well suited to recording fine urban structures and vegetation profiles and has potential uses in remote sensing and urban planning. Better results are assured with more training and novel loss functions or architectural design. This work identifies the potential of deep learning super resolution techniques to enhance satellite imagery to make remote sensing applications more efficient. Additional research will further develop the model to be more advanced and utilize it on other real-world datasets to make it more practically relevant in satellite image analysis.

REFERENCES

[1] 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.
[2] Y. Li, Y. Zhang and Z. Zhu, "Learning Deep Networks under Noisy Labels for Remote Sensing Image Scene Classification," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 3025-3028, doi: 10.1109/IGARSS.2019.8900497.
[3] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
[4] Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2017, pp. 136-144
[5] S. Yan and M. Yao, "Dense Residual Network for Text Image Super-Resolution," 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2024, pp. 1172-
[6] 1176, doi: 10.1109/CVIDL62147.2024.10603757. [6] Zhu, H., & Sun, W. (2020). Super-Resolution for Satellite Imagery Using Dense Skip Connections and Residual Learning. In Remote Sensing, 12(8), 1267
[7] Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image Super-Resolution Using Deep Convolutional Networks. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2), 295-307.
[8] I. Goodfellow et al., "Generative Adversarial Nets," Advances in Neural Information Processing Systems, 2014.
[9] J. Kim, J. K. Lee, and K. M. Lee, “Accurate Image Super-Resolution Using Very Deep Convolutional Networks,” CVPR 2016, pp. 1646-1654.
[10] X. Wang et al., "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," ECCV Workshops, 2018.
[11] X. Wang et al., "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," ICCVW, 2021.
[12] A. Lugmayr, M. Danelljan, and R. Timofte, “SRFlow: Learning the Super-Resolution Space with Normalizing Flow,” ECCV 2020, pp. 715–732.

