**ABSTRACT**

Image Super-Resolution (ISR) aims to reconstruct high-resolution (HR) images from low-resolution (LR) inputs, enhancing image clarity and detail. In this project, we implement a Super-Resolution Generative Adversarial Network (SRGAN), a deep learning-based model that generates photorealistic HR images with perceptual quality. SRGAN leverages a generator-discriminator architecture, where the generator learns to upscale images, and the discriminator distinguishes between real HR images and generated ones. Additionally, a perceptual loss function based on high-level feature representations from a pre-trained VGG network ensures sharper and more visually pleasing results.

Our model is trained on a large dataset of natural images and is capable of recovering fine textures and details that are typically lost in traditional interpolation or CNN-based approaches. The integration of adversarial training encourages the generator to produce outputs that are perceptually closer to real-world images. This approach significantly outperforms conventional techniques in terms of visual fidelity and structural similarity, making it highly suitable for applications in medical imaging, satellite photography, security surveillance, and digital media enhancement.

**INTRODUCTION**

In today’s digital world, high-resolution (HR) images are crucial for various applications such as medical diagnostics, satellite imaging, surveillance, and media production. However, obtaining HR images is often expensive, time-consuming, or technically challenging due to hardware limitations. This has led to the development of Image Super-Resolution (ISR) techniques, which aim to reconstruct HR images from low-resolution (LR) inputs. Traditional methods, such as bicubic interpolation, suffer from poor performance and fail to recover high-frequency details, resulting in overly smooth and unrealistic outputs.

When diagnosing and analyzing illnesses, high-resolution medical photographs are crucial. High-Resolution images have f iner details, and essential features make it easier to analyze images and make an accurate diagnosis. Common models like MRI, CT-Scan, and X-RAY have resolution constraints due to factors such as motion limits, dose limitations, and hardware costs.

Low-resolution images make it difficult to analyze the images for diagnosis. low-resolution images adversely affect the identification of crucial features. This leads to the im proper diagnosis. The goal of super-resolution techniques is to replicate the high-resolution image using the low-resolution image, while preserving the finer details, essential features. Traditional interpolation methods often fail to preserve these f iner details.

Low-resolution photographs can be converted into realistic, high-resolution images using a generative adversarial network known as SRGAN (Super-resolution Generative Adversarial Networks). It is made up of a generator and a discriminator. The discriminator makes a distinction between photos with low and high resolution. while the generator network creates the images. Adversarially training the discriminator and generator produces perceptually realistic visuals. [1].

Despite the proven potential of SRGAN in enhancing natural images, its application in the medical imaging domain is very unexplored. There is a lack of comprehensive studies addressing extreme upscaling in medical imaging using deep learning-based super-resolution methods. we apply SRGAN to upscaling medical images from low-resolution to high resolution images

* When upscaling medical images from low-resolution to high resolution, we use SRGAN
* We employ the PSNR and SSIM methodologies to assess the outcomes

**LITERATURE REVIEW**

Image Super-Resolution Using Deep Convolutional Networks by Dong, Chao, et al, the author proposed a deep learning method for single image super-resolution (SR). their method directly learns an end to end mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolutional network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers [1].

Khaledyan, Donya, et al. proposed a low-cost implementation of bilinear and bicubic image interpolation aimed at achieving real-time image super-resolution, especially optimized for mobile applications. Recognizing that image interpolation plays a crucial role in super-resolution algorithms, the authors designed two hardware-efficient methods that are both computationally light and effective. Their experiments on both synthetic and real-world image sequences demonstrated that the proposed schemes not only perform efficiently but also maintain acceptable output quality. This work validates the potential of classic interpolation techniques when optimized for hardware constraints, although such methods still fall short in reconstructing fine textures and high-frequency details when compared to learning-based and GAN-based methods like SRGAN [2].

Kim et al. (2016) proposed VDSR (Very Deep Super-Resolution), a deep learning model using a 20-layer CNN to improve image super-resolution accuracy. They introduced residual learning to help the model learn the difference between low-resolution and high-resolution images, making training more efficient and faster. VDSR significantly improved PSNR and SSIM scores compared to earlier methods like SRCNN. It also allowed for flexible scaling factors with a single model. However, while it produced accurate images, it often lacked fine texture details, which later GAN-based models like SRGAN addressed for more perceptual realism [3].

Wang et al. (2018) introduced ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), an improved version of SRGAN that significantly enhanced the perceptual quality of generated images. They proposed a new architecture using Residual-in-Residual Dense Blocks (RRDB), which improved feature representation without relying on batch normalization. ESRGAN also replaced the original perceptual loss with a more accurate loss computed using the VGG network before activation layers, capturing finer textures. The model produced sharper and more realistic details compared to SRGAN, setting a new standard in perceptual image super-resolution. ESRGAN has since been widely adopted in both academic and practical image enhancement tasks [4].

Zhang, Yulun, et al. Image super-resolution using very deep residual channel attention networksauthor  observe that deeper networks for image SR are more difficult to train. The low-resolution (LR) inputs and features contain abundant low-frequency information, which is treated equally across channels, hence hindering the representational ability of CNNs. To solve these problems, we propose the very deep residual channel attention networks (RCAN). Specifically, we propose residual in residual (RIR) structure to form very deep network, which consists of several residual groups with long skip connections. Each residual group contains some residual blocks with short skip connections. Meanwhile, RIR allows abundant low-frequency information to be bypassed through multiple skip connections, making the main network focus on learning high-frequency information [5].

Ledig et al. (2017) introduced SRGAN, the first framework to leverage Generative Adversarial Networks for single-image super-resolution. The model combines a deep residual network generator with a discriminator trained to distinguish high-resolution images from generated ones. A key innovation was the use of perceptual loss, based on the feature space of a pre-trained VGG network, which encourages photo-realistic and texture-rich outputs. This allowed SRGAN to generate sharper and more visually appealing images compared to traditional pixel-wise approaches. The paper set a new benchmark in perceptual quality and inspired a wave of research in GAN-based super-resolution techniques [6].

**METHODOLOGY**

Our work explores the use of improved Super-Resolution Generative Adversarial Networks (SRGAN) for enhancing medical imaging quality. The work draws on a hand-picked dataset of 5,000 medical scans, artificially down sampled to 32×32 pixel resolution to simulate the difficult clinical imaging scenarios. Our main objective is to upscale these images from 32×32 resolution to 128×128 resolution while preserving diagnostically important characteristics.

*A. Image Generation Network*

The core of the improvement system is a very sophisticated neural network designed especially for medical image super resolution. The structure starts processing with an initial convolutional layer with Parametric ReLU activation, chosen for its adaptive learning characteristics. The network’s central processing component incorporates multiple specialized residual modules, each containing convolutional operations, normalization layers, and strategic skip pathways. These elements work synergistically to preserve crucial anatomical structures and subtle tissue patterns.

For resolution enhancement, we implement advanced sub pixel convolutional operations (commonly referred to as pixel shuffling), a technique that progressively increases image dimensions while minimizing reconstruction artifacts. The final processing stage outputs refined 128×128 resolution images suitable for clinical evaluation. [1]

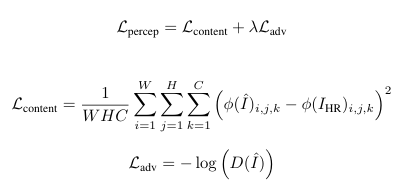
*B. Discriminative Evaluation Network*

Complementing the generation system, we developed a sophisticated verification network that assesses image authen ticity. This component functions as a specialized classifier, dis tinguishing between genuine high-resolution medical images and enhanced versions. Its architecture features a progressive series of convolutional layers with carefully tuned filter di mensions, integrated with LeakyReLU activation functions for nuanced feature detection. The terminal layers employ dense connections to generate authenticity probabilities, providing critical feedback for system improvement. [1]

*C. Loss Functions- Balancing Quality and Realism*

The loss function of SRGAN is called as a perceptual loss function which is calculated by weighted sum of two components: content loss (also known as VGG loss) and adversarial loss. [1]:

* **Content Loss:** The feature maps of the produced super resolution image and the high-resolution real image are used to compute this loss. Specifically, it calculates the mean squared error between two images’ VGG feature maps. This aids the generator in creating visuals that are more realistic and perceptually comparable.
* **Adversarial Loss:** This loss, which comes from the GAN discriminator loss, is employed to make the generator create very indistinguishable images. The discriminator has been trained to distinguish between produced and genuine images.



By using this perceptual loss function SRGAN aims to produce super-resolution images

*D. Training Protocol*

We trained the model on 5000 image pairs from a Pneumonia X-RAY image dataset. We pre-processed the images into (32 × 32) and (128 × 128) pairs. The generator is tasked with tricking the discriminator, and the training process alternates between updating the discriminator and the generator.

*E. Model Architecture*

Super-Resolution Generative Adversarial Network (SRGAN) is a deep learning model designed to generate high-resolution images from low-resolution inputs. It consists of two main components:

* **Generator**: A deep residual network that learns to up sample low-resolution images into high-resolution ones. It uses residual blocks with skip connections to preserve image details and stability during training.
* **Discriminator**: A convolutional neural network that tries to distinguish between real high-resolution images and the ones generated by the generator. It pushes the generator to produce more realistic outputs.

SRGAN is trained using a combination of **adversarial loss** (from GAN) and **content loss** (typically using VGG-based perceptual loss), enabling it to generate sharper and more visually appealing images compared to traditional methods.

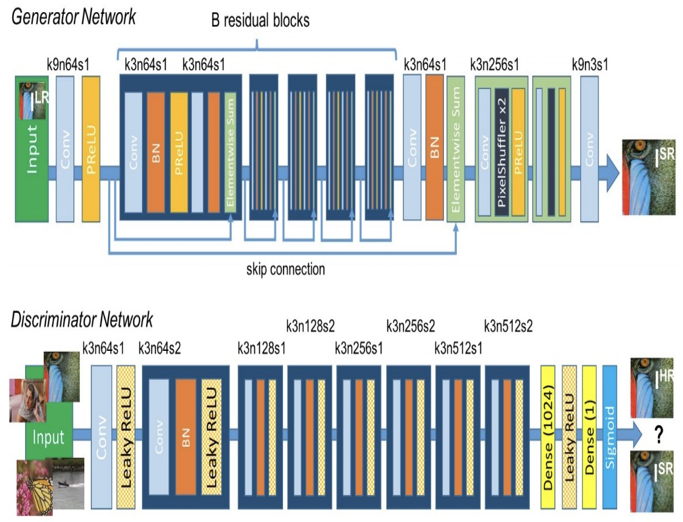


Figure 1 Architecture of SRGAN

*F. Website Integration*

The SRGAN project utilizes a modular full-stack architecture to deliver a seamless user experience and efficient image processing. At the heart of the web interface is Next.js, a modern React-based framework that excels in performance and server-side rendering. This ensures a responsive and interactive platform, allowing users to effortlessly upload low-resolution images and instantly view their corresponding high-resolution outputs, enhancing user engagement and satisfaction.

On the backend, Flask is employed to facilitate communication between the frontend and the machine learning model. When a user submits an image, it is transmitted to the server as form data. The Flask application receives this data, performs necessary preprocessing tasks, and then forwards the image to the SRGANmodel for processing. Once the model has generated the enhanced high-resolution image, the server sends the result back to the frontend, where it is displayed for the user to review.

The SRGAN model itself is built using Python and leverages cutting-edge deeplearning techniques to enhance image resolution. The architecture consists of two key components: a generatornetwork responsible for creating high-quality, upscaled images, and a discriminatornetwork that ensures the generated outputs appear perceptually realistic and aligned with high-resolution images.

This combination of Next.js, Flask, and Python results in a highly functional, real-world deployment scenario. The design not only makes the SRGAN model accessible to users but also offers an intuitive and visually engaging web interface that simplifies complex deep learning tasks. By merging these technologies, the project exemplifies how advanced AI techniques can be effectively integrated into practical, user-friendly applications.

**Key Features:**

* **Real‑time Upload & Preview**: Instantly upload and view super‑resolved images.
* **High‑Quality Enhancements**: State‑of‑the‑art SRGAN generator for sharper, more detailed outputs.
* **Perceptual Realism**: Discriminator‑driven training ensures results closely match true high‑resolution images.
* **Scalable Architecture**: Modular frontend (Next.js) and backend (Flask) enable easy updates and containerized deployment.
* **User‑Friendly Interface**: Clean, responsive UI designed for intuitive image comparison and analysis.

The website architecture for the SRGAN project follows a client-server model, designed for seamless integration of machine learning with a web interface. The frontend, built with Next.js, allows users to upload low-resolution images and view enhanced results in real-time. On the backend, a Flask server handles requests, forwards images to the SRGAN model (implemented in Python), and returns the super-resolved output. This modular design ensures fast, scalable, and responsive image processing, simulating real-world deployment of AI applications.

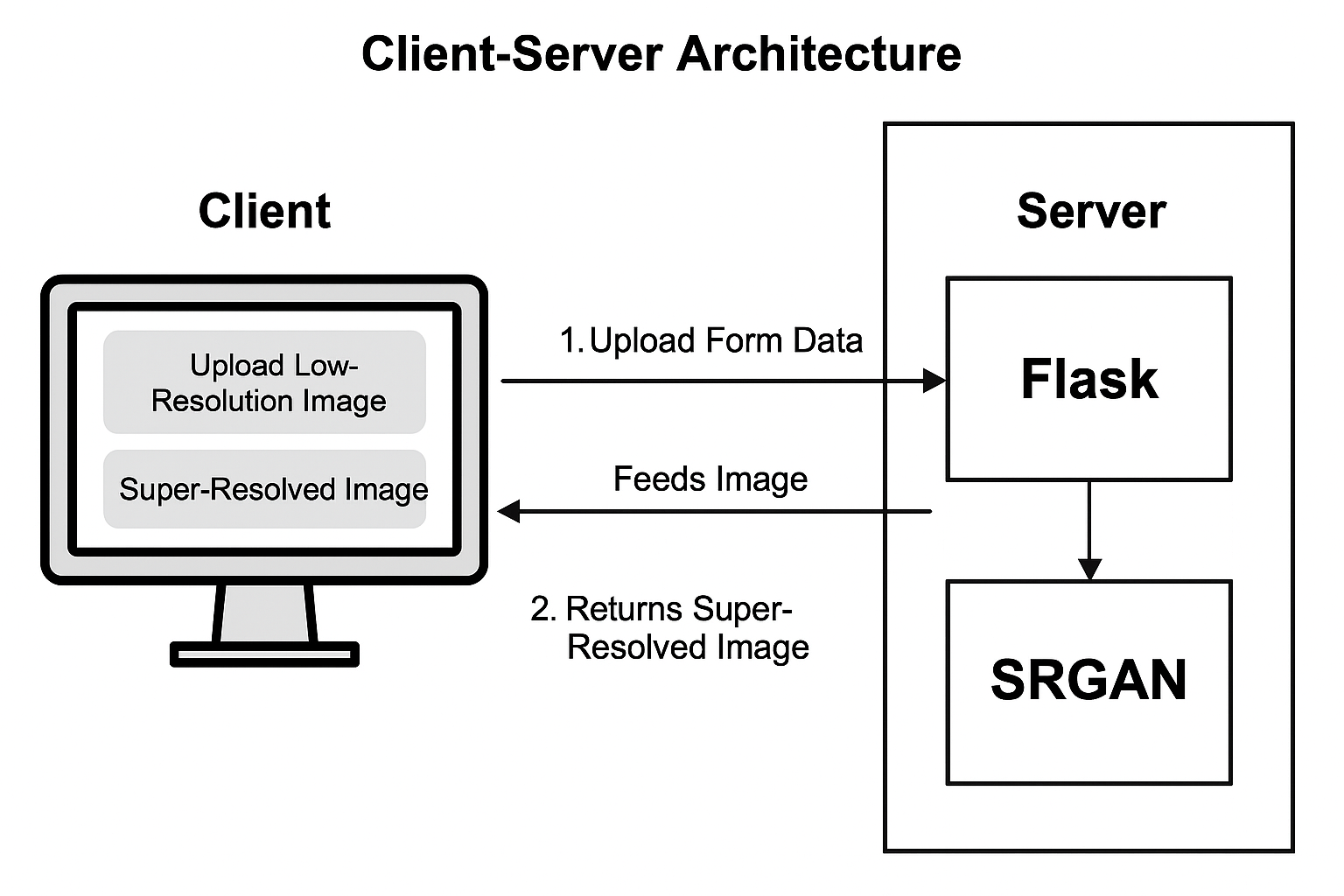
****

Figure 2 Architecture of the Website

*G. System Design and User flow*

The system is based on a Client-Server Architecture, where the user-facing frontend and the backend model processing logic are cleanly separated, enabling scalability and ease of maintenance.

On the client side, users interact with the application through a user-friendly interface built with Next.js. When a user uploads a low-resolution image, the image is packaged as form data and sent to the backend using an HTTP POST request.

The server side is powered by a Flask backend, which acts as the intermediary between the client and the SRGAN model. Upon receiving the uploaded image, the server performs preprocessing steps required by the model and feeds the image into the SRGAN pipeline. The SRGAN model then performs super-resolution, generating a high-quality, upscaled version of the image.

Once the processing is complete, the Flask API sends the resulting high-resolution image back to the frontend. The client then receives and displays both the original and the super-resolved images side by side, offering users an intuitive and visually interactive comparison. This flow effectively demonstrates the impact of the SRGAN model in a real-time web environment.

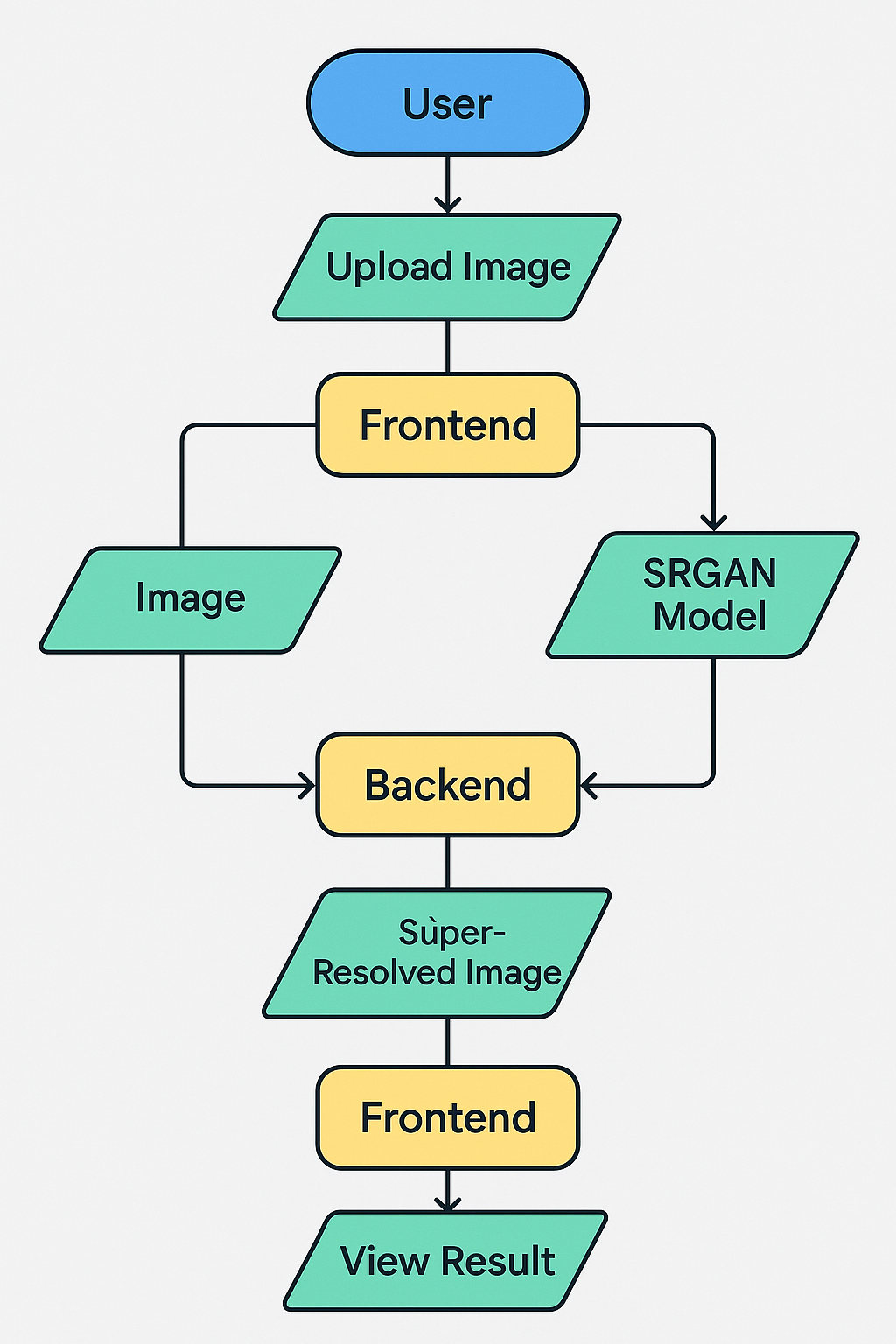
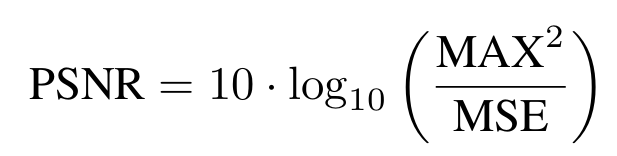


Figure 3 User flow diagram

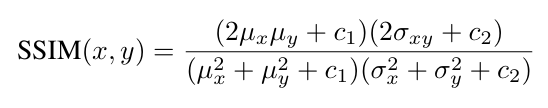
**EVALUATION METRICS**

We used a set of assessment measures to gauge the effective ness of our SRGAN model, which allowed us to ascertain both the model’s overall efficiency and the caliber of the images it generated. These indicators provide us with a comprehensive picture of the model’s effectiveness across many scenarios.

* **Peak Signal-to-Noise Ratio (PSNR):** This image fidelity number contrasts the highest possible signal strength with the degrees of reconstruction noise. Greater PSNR values (such as 30–50 dB for radiological pictures) in medical imaging usage signify better preservation of diagnostic information. The logarithm calculation in decibels is:



* **Structural Similarity Index (SSIM):** With an emphasis on structural elements like textures and edges, SSIM calculates the perceived similarity between the generated image and the ground truth. In contrast to PSNR, SSIM assesses the image according to how the human eye interprets it. The SSIM is calculated as follows:



where σ2 x and σ2 y are the pictures’ variances, σxy is the covariance, and µx and µy are their average pixel values. In low-value situations, the division is stabilized by the constants c1 and c2.

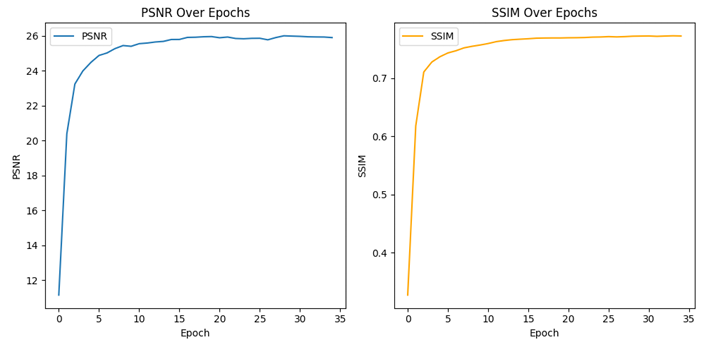
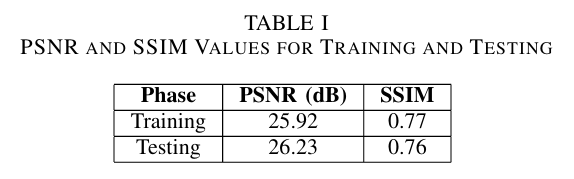


Figure 4 PSNR and SSIM vs Epochs



* **Learning Curves:** We also monitor the learning curves of the model during training, which show how the con tent and adversarial losses evolve. These curves provide insights into how well the model is converging over time, helping to ensure the training process is stable and efficient.

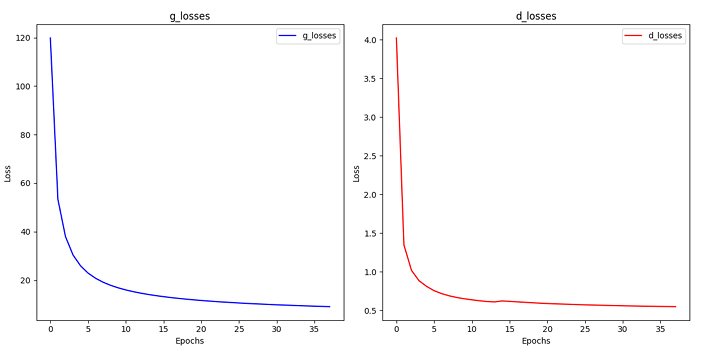


Figure 5 Generator loss and Discriminator loss

**RESULTS AND DISCUSSION**

The proposed SRGAN model, trained on a medical image dataset, achieved impressive performance in turning low resolution inputs into high-resolution photos. During training, the model reached a Peak Signal-to-Noise Ratio (PSNR) of 25.2 and a Structural Similarity Index (SSIM) of 0.77, demonstrating its ability to generate perceptually accurate images. In testing, the model’s PSNR increased to 26.23, with an SSIM of 0.76, indicating that it successfully generalizes to unseen data. The generator loss during training was 9.848, while the discriminator loss was 0.556, showing that the adversarial training process was effectively balancing the generation of realistic images and the discriminator’s ability to differentiate between real and fake images. This model demonstrates strong potential for applications in medical image enhancement, where preserving critical features and improving image qual ity is essential for accurate diagnoses. Its promising results highlight the robustness and generalizability of the SRGAN model, with minimal overfitting observed.

*A. Visual Results*

The LR input photos are compared side by side in Figure 3, the SR images produced by the SRGAN model, and the HR ground truth images. As shown in the figure, the SR images closely resemble the HR images, with the SRGAN model effectively recovering fine details and textures that were lost in the LR images. This highlights the ability of SRGAN to restore meaningful features that are important for medical image analysis.

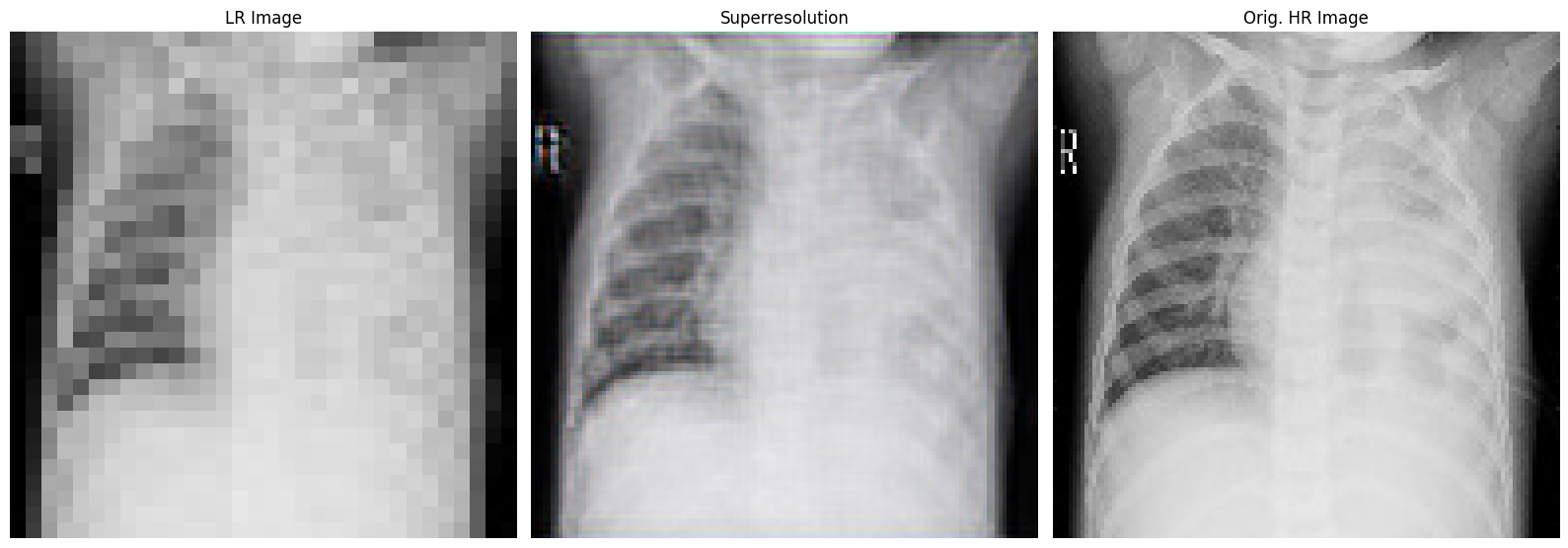
****

Figure 6 Comparison of Low-Resolution (LR), Super-Resolved (SR), and High

*B. Visual Inspection*

From a visual standpoint, the SRGAN model has demonstrated the capability to generate high-resolution images that preserve crucial structural elements, such as edges and textures. These results suggest that the model is not only effective in enhancing the resolution but also in maintaining key diagnostic features, making it suitable for medical image applications where preserving fine details is critical.

**CONCLUSION AND FUTURE WORK**

This study highlights the remarkable effectiveness of SRGAN in enhancing the resolution of medical images—an essential task where image clarity can directly impact clinical decisions. During evaluation, the model achieved a PSNR of 25.2 and SSIM of 0.77 on the training dataset, and slightly better results on the testing dataset with a PSNR of 26.23 and SSIM of 0.76. These metrics indicate the model’s ability to produce outputs that are not only sharp and detailed but also closely aligned with the ground truth high-resolution images.

Beyond the numbers, what truly stands out is the SRGAN model’s capacity to retain critical structural and textural features in medical images. These details are vital for accurate diagnosis and clinical interpretation, especially in areas like radiology or histopathology where subtle features can make a significant difference.

Furthermore, the relatively low generator and discriminator losses provide additional evidence of the model’s stability and realism in generated outputs. The consistency between predicted and actual images suggests that SRGAN effectively learns the underlying distribution of medical image data, enabling it to generate perceptually convincing results.

In conclusion, SRGAN proves to be a powerful tool for medical image super-resolution. By delivering high-quality, realistic enhancements, it supports better visualization, potentially improvingdiagnosticaccuracy and promoting better patient outcomes. Its integration into clinical workflows could pave the way for more accessible and reliable imaging, especially in settings where high-resolution equipment may be limited.

Looking forward, several improvements can be explored. Incorporating attention mechanisms could help the model better focus on subtle, clinically important features. Expanding the training to larger and more diverse datasets could improve its robustness and generalization across different medical imaging modalities. Further, experimenting with domain specific perceptual loss functions may enhance fine-grained detail reconstruction. Lastly, optimizing the model for faster inference and validating its performance in real-world clinical environments would move this research closer to practical application.

**ACKNOWLEDGMENT**

Sincere thanks are extended by the authors to the Chaitanya Bharathi Institute of Technology for providing the infrastructure and the resources required to finish this research successfully. Special thanks to the Department of Artificial Intelligence and Machine Learning for granting access to high performance computing systems, which played a crucial role in the efficient execution of our experiments. The advanced computational facilities and resources available were integral to the progress and results achieved in this research. We deeply appreciate the support provided, which enabled the successful development and implementation of our SRGAN model for medical image enhancement.

**REFERENCES**

[1] C. Ledig, L. Theis, F. Husz´ ar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” arXiv preprint arXiv:1609.04802, 2017.

[2] D. Khaledyan et al., “Low-cost implementation of bilinear and bicubic image interpolation for real-time image super-resolution,” in 2020 IEEE Global Humanitarian Tec1hnology Conference (GHTC), pp. 1–7, IEEE, 2020.

[3] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in Proceedings of the IEEE conference on 1computer vision and pattern recognition, pp. 1646–1654, 2016.

[4] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, “Esrgan: Enhanced super-resolution generative adversarial networks,” in Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018.

[5] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual channel attention networks,” in Proceedings of the European Conference on Computer Vision (ECCV), pp. 286–301, 2018