

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING**

**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY**

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana

[www.cbit.ac.in](http://www.cbit.ac.in)

**A Project Report on**

**Image Super-Resolution using-SRGAN**

Submitted in partial fulfilment of the requirements for the award of degree

**BACHELOR OF ENGINEERING**

**in**

**ARTIFICIAL INTELLIGENCE & MACHINE LEARNING**

**by**

**ABHINAY SABHANAM 160122729021**

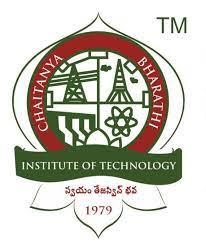
**BERA KRISHNA CHAITANYA 160122729026**

**KUNTIGORLA RAMESH 160122729301**

Under the guidance of

**Ms. FALAK NAAZ**

**Assistant Professor AIML Dept.**



**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY**

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana

[www.cbit.ac.in](http://www.cbit.ac.in)

**CERTIFICATE**

This is to certify that the project titled “**Image Super-Resolution using-SRGAN**” is the bonafide work carried out by  **ABHINAY SABHANAM (160122729021), BERA KRISHNA CHAITANYA (160122729026) and KUNTIGORLA RAMESH (160122729301),** students of B.E. (AIML) of Chaitanya Bharathi Institute of Technology(A), Hyderabad, affiliated to Osmania University, Hyderabad, Telangana(India) during the academic year 2024-2025, submitted in partial fulfilment of the requirements for the award of the degree in **Bachelor of Engineering Artificial Intelligence & Machine Learning** and that the project has not formed the basis for the award previously of any other degree, diploma, fellowship or any other similar title.

**Supervisor Incharge of Mini Project Head of the Department**

**Ms. Falak Naaz**  **Mr. Panduraju Pagidimalla** **Dr. Y. Rama Devi**

**Asst. Professor ,AI&ML Asst. Professor ,AI&ML Professor, Head of Dept. ofAI&ML**

**Place:** CBIT-Hyderabad

**Date:**

****

**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY**

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana www.cbit.ac.in

**Vision of Institute**

To be the Centre of Excellence in Technical Education and Research.

**Mission of Institute**

To address the Emerging needs through Quality Technical Education and Advanced Research.

**Quality Policy**

CBIT imparts value based Technical Education and Training to meet the requirements of students, Industry, Trade/ Profession, Research and Development Organizations for Self-sustained growth of Society.



**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY** (An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana www.cbit.ac.in

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Vision of the Department**

To produce professionals in artificial intelligence and machine learning through the best possible education, acquire international recognition as a destination, and advance society in exciting and creative ways.

**Mission of the Department**

• Impart rigorous training to generate knowledge through the state-of-the-art concepts and technologies in Artificial Intelligence and Machine Learning.

• Develop technical proficiency in students through creativity and leadership.

• Encourage lifelong learning, social responsibility, environmental conservation, and professional ethics.

• Establish centres of excellence in leading areas of computer and artificial intelligence disciplines.

**Program Educational Objectives (PEOs)**

**PEO1:** Using a solid foundation in mathematical, scientific, engineering, and current computing principles, formulate, analyse, and resolve engineering issues.

**PEO2:** Apply artificial intelligence theory and concepts to analyse the requirements, realise technical specifications, and design engineering solutions.

**PEO3:** Through cross-disciplinary projects and a variety of professional activities, demonstrate technical proficiency, AI competency, and foster collaborative learning and a sense of teamwork.

**PEO4:** Provide graduates with solid knowledge, competence, and soft skills that will enable them to ethically contribute to societal demands and achieve sustainable advancement in emerging computer technologies through lifelong learning.



**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY** (An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana www.cbit.ac.in

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Program Outcomes (POs)**

**PO1.Engineering Knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization for the solution of complex engineering problems

**PO2.Problem analysis:** Identify, formulate, research literature, and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

**PO3.Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety, and cultural, societal, and environmental considerations.

**PO4.Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

**PO5.Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools, including prediction and modelling to complex engineering activities, with an understanding of the limitations.

**PO6.The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

**PO7.Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

**PO8.Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice. Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

**PO9.Cosmmunication:** Communicate effectively on complex engineering activities with the engineering community and with the society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

**PO10.Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

**PO11.Life-long learning:** Recognise the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

**Program Specific Outcomes (PSOs)**

**PSO1.** Apply the principal concepts of AI Engineering to design, develop, deploy and prototype AI Subsystems.

**PSO2.** Apply the knowledge gained pertaining to data storage, data analytics and AI concepts to solve real world business problems.

**PSO3**. Apply, analyse, design, develop, and test principles of AI concepts on Intelligent Systems.



**CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY** (An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,

Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana www.cbit.ac.in

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**MINI PROJECT (22AMC12)**

**Course Objectives:** The aim of course is

1. To explore the literature and formulate a project proposal.
2. To enhance presentation skills and technical writing proficiency.
3. To provide solutions by using modern tools.
4. To Expose Students to Project Based Learning.
5. To effective presentation and documentation.

**Course Outcomes:** After completion of this course, students will be able to

1. Interpret Literature the purpose of formulating a project proposal.
2. Plan, Analyze, Design and implement a project.
3. Find the solution of an identified problem with the help of modern Technology and give priority to real time scenarios.
4. Plan to work as a team and to focus on getting a working project done and submit a report within a stipulated period of time.
5. Prepare and submit the Report and deliver a presentation.

**CO-PO/PSO Articulation Matrix:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 | PO11 | PO12 | PSO1 | PSO2 | PSO3 |
| CO1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CO2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CO3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CO4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CO5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**DECLARATION**

We hereby declare that the project entitled “**Image Super-Resolution using-SRGAN**” submitted for the B.E (CAIML) degree is our original work and the project has not formed the basis for the award of any other degree, diploma, fellowship or any other similar titles.

**Names and Signatures of the Students**

**ABHINAY SABHANAM**

**(160122729021)**

**BERA KRISHNA CHAITANYA**

**(160122729026)**

**KUNTIGORLA RAMESH**

**(160122729301)**

**Place:** CBIT-Hyderabad

**Date:**

**ACKNOWLEDGEMENT**

The satisfaction that accompanies the successful completion of the task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crown all the efforts with success.

We show gratitude to our honourable Principal **Dr. C. V. Narasimhulu**, for providing all facilities and support.

We are particularly thankful for **Dr. Y. Rama Devi**, the Head of the Department, Department of Artificial Intelligence and Machine Learning, her guidance, intense support, and encouragement, which helped us to mould our project into a successful one.

We wish to express our deep sense of gratitude to out supervisor **Ms. Falak Naaz,** Assistant Professor, Department of Artificial Intelligence and Machine Learning, for her able guidance and useful suggestions, which helped us in completing the project in time.

We are particularly thankful to our Project Coordinators **Mr. Panduraju Pagidimalla,** Assistant Professorand **Dr. Prabhakar Kandukuri** Professor, Department of Artificial Intelligence and Machine Learning, their guidance, intense support, and encouragement, which helped us to mould our project into a successful one.

We also thank all the staff members of the AIML department for their valuable support and generous advice.

**Names and Signatures of the Students**

**ABHINAY SABHANAM (160122729021)**

**BERA KRISHNA CHAITANYA (160122729026)**

**KUNTIGORLA RAMESH (160122729301)**

**Place:** CBIT-Hyderabad

**Date:**

**CONTENTS**

ABSTRACT ……………………………………………………………………………….1

1.INTRODUCTION ………………………………………………………………………2

1.1 Problem Definition…………………………………………………………3

1.2 Methodology…………………………………………………………………3

1.3 Outline of the result……………………………………………………….4

1.4 Scope of the project………………………………………………………..4

1.5 Organization of the report………………………………………………..4

2.LITERATURE REVIEW ………………………………………………………………5

2.1 Introduction to Domain Technology……………………………………5

2.2 Existing Methods…………………………………………………………..5

2.3 Related Works………………………………………………………………6

2.4 Tools and Technologies used …………………………………………...8

3.DESIGN OF PROPOSEDSYSTEM…………………………………………………10

3.1 Image generation Network …………..………………………………….10

3.2 Discriminative Evaluative Network…………………………………….10

3.3 Loss Functions- Balancing Quality and Realism……………………11

3.4 Training Protocol.…………………………………………………………..11

3.5 Model Architecture ………………………………………………………..11

3.6 Website Integration ……………………………………………………….13

3.7 System Design and User flow …………………………………………..14

4.IMPLEMENTATION OF THE PROPOSED SYSTEM …………………………..16

4.1 Module Description ……………………………………………………….16

4.2 Algorithms …………………………………………………………………..17

4.3 Dataset description ……………………………………………………….18

4.4 Testing Process …………………………………………………………….19

5.RESULTS AND DISCUSSION……………………………………………………….21

6.CONCLUSION AND FUTURE WORK…………………………………………..….23

7.REFERENCES…………………………………………………….……………………24

8.SUBMISSION OF RESEARCH PAPER…………………………………………….25

**LIST OF FIGURES**

FIGURE 1………………………………………………. ARCHITECTURE OF SRGAN

FIGURE 2………………………… …………..ARCHITECTURE OF THE WEBSITE

FIGURE 3………………………………………………………USER FLOW DIAGRAM

FIGURE 4………………………………………………….PSNR & SSIM VS EPOCHS

FIGURE 5…………………………………………………………….LOSS FUNCTIONS

FIGURE 6……………..……………COMPARSION IN RESOLUTION OF IMAGES

**LIST OF TABLES**

TABLE1………………PSNR AND SSIM VALUES FOR TRAIN AND TEST SETS

**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.

* + - 1. **INTRODUCTION**

**1.1 Problem Definition**

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.

[1]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.

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

**1.2 Methodology**

The SRGAN architecture consists of two main components: a generator and a discriminator.

* **Generator**: This is a deep residual network (ResNet) that takes a low-resolution image as input and attempts to generate a super-resolved version. It uses residual blocks, skip connections, and upsampling layers to refine image quality and maintain spatial details.
* **Discriminator**: This network acts as a binary classifier that distinguishes between real high-resolution images and those generated by the generator. It guides the generator to produce more realistic images.
* **Loss Functions**:
  + **Content Loss**: Instead of using pixel-wise MSE alone, SRGAN uses perceptual loss computed from feature maps of a pre-trained VGG network to maintain texture fidelity.
  + **Adversarial Loss**: Encourages the generator to produce outputs that are indistinguishable from real images.
  + The final loss is a weighted combination of content and adversarial losses.

The model is trained end-to-end using a dataset of paired low- and high-resolution images. During training, the generator learns to enhance image details, while the discriminator ensures that the generated images are perceptually closer to real images.

**1.3 Outline of the Result**

This report is structured into six chapters to comprehensively document the development of the Image Super-Resolution project using SRGAN. The first chapter introduces the project by outlining the problem statement, its relevance in the field of computer vision, the key objectives, methodology adopted, scope, and the structure of the report. The second chapter provides a literature review of existing image enhancement techniques, traditional upscaling methods, and recent advancements using deep learning, particularly GAN-based approaches, highlighting the motivation for using SRGAN.

Chapter three explains the system design, detailing the architecture of the SRGAN model, including the generator, discriminator, loss functions, and the workflow from input to output. The fourth chapter focuses on the implementation, describing dataset preparation, model training, hyperparameter tuning, and the development environment used. Chapter five presents the results and evaluation, showcasing qualitative outputs, quantitative metrics such as PSNR and SSIM, and a comparison with baseline methods. The final chapter summarizes the project outcomes, key learnings, and outlines possible directions for future improvements and real-world deployment. The report concludes with a list of references comprising academic papers, datasets, tools, and libraries consulted during the project.

**1.4 Scope of the Project**

This project is focused on implementing and evaluating an Image Super-Resolution model based on SRGAN. It aims to improve the resolution and perceptual quality of low-resolution images. The scope includes model training, testing on benchmark datasets (e.g., Set5, Set14), and evaluating performance using both qualitative and quantitative metrics. However, the project does not extend to multi-frame video super-resolution or real-time deployment on mobile devices.

**1.5 Organization Report**

This report is organized into six chapters covering the development of the Image Super-Resolution project using SRGAN. Chapter one introduces the problem, objectives, scope, and methodology. Chapter two reviews related work, traditional upscaling methods, and recent deep learning approaches, highlighting the relevance of SRGAN. Chapter three details the system design, including model architecture and workflow. Chapter four explains implementation aspects like dataset preparation, training, and tools used. Chapter five presents results, evaluation metrics (PSNR, SSIM), and comparisons. Chapter six concludes with key outcomes and future enhancements, followed by references.

**2. LITERATURE REVIEW**

**2.1 Introduction to Problem Domain and Technology**

Image Super-Resolution (ISR) is a key area in computer vision that focuses on reconstructing high-resolution (HR) images from low-resolution (LR) inputs. It has wide applications in medical imaging, satellite imagery, surveillance, and more, where visual detail is crucial. Traditional interpolation methods like bicubic scaling often produce blurry results lacking fine detail.

Recent advances in deep learning, particularly Generative Adversarial Networks (GANs), have significantly improved the quality of super-resolved images. SRGAN (Super-Resolution Generative Adversarial Network) is a pioneering model that leverages the power of GANs to generate photo-realistic high-resolution images. It uses a generator-discriminator setup and incorporates perceptual loss based on high-level feature representations, enabling more accurate texture and detail restoration compared to conventional approaches.

**2.2 Existing Methods**

**Bicubic Interpolation** is a traditional method that estimates pixel values using weighted averages of nearby pixels. It’s fast and easy to implement but often results in blurry images with weak edges.  
*Performance:* PSNR ~23–25 dB | SSIM ~0.60–0.65  
*Summary:* Simple and fast, but poor at preserving texture.

**SRCNN** [1]was one of the first deep learning models for image super-resolution. It used a shallow 3-layer convolutional network to map low-resolution images to high-resolution ones.  
*Performance:* PSNR ~27.5 dB | SSIM ~0.89 (on Set5)  
*Summary:* A major step forward, though limited by its shallow depth.

**VDSR** [2] introduced a deeper 20-layer CNN along with residual learning to improve performance and training. It could handle multiple scales in one model.  
*Performance:* PSNR ~28.8 dB | SSIM ~0.91 (on Set5)  
*Summary:* Produced sharper results, but still lacked natural textures.

**FSRCNN** [3] was designed to be a faster variant of SRCNN using deconvolution layers and smaller filters.  
*Performance:* PSNR ~27.6 dB (on Set5)  
*Summary:* Efficient in speed, especially on mobile devices, but not as accurate as deeper models.

**SRGAN** [4] was a breakthrough that introduced GANs into the SR domain. It used a generator-discriminator setup and perceptual loss from a VGG network to create photo-realistic images.  
*Performance:* PSNR ~29.4 dB | SSIM ~0.91 (on Set5)  
*Summary:* Visually impressive, though slightly lower on traditional accuracy metrics.

**ESRGAN** [5] built on SRGAN by using Residual-in-Residual Dense Blocks (RRDB) and improved loss functions. It generated even sharper and more realistic results.  
*Performance:* PSNR ~29.7 dB | SSIM: ~0.92  
*Summary:* Great balance between texture quality and accuracy.

**RCAN** [6] used channel attention mechanisms and a deep residual-in-residual structure to focus on important features.  
*Performance:* PSNR ~30.6 dB | SSIM ~0.92 (on Set5)  
*Summary:* Very accurate but computationally demanding.

**LapSRN** [7] adopted a Laplacian pyramid approach to reconstruct images in a coarse-to-fine manner.  
*Performance:* PSNR ~28.8 dB | SSIM ~0.91  
*Summary:* Balanced speed and quality well.

**EDSR** [8] removed batch normalization and used deeper residual blocks, leading to one of the highest PSNR values at the time.  
*Performance:* PSNR ~32.6 dB | SSIM ~0.94 (on Set5)  
*Summary:* Extremely accurate but less visually realistic than GAN-based methods.

**Bilinear Interpolation** [9] focused on optimizing classic interpolation for embedded systems. Though not ideal for quality, it’s efficient in resource-constrained environments.  
*Performance:* PSNR ~24–25 dB | SSIM ~0.65  
*Summary:* Fast and hardware-friendly, but lacks detail restoration.

**2.3 Related Works**

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].

**2.4 Tools/Technologies Used in SRGAN**

1. **Deep Learning Frameworks**:
   * **TensorFlow** or **PyTorch**: These are the primary frameworks used to implement SRGAN. They provide powerful APIs for building and training deep neural networks.
   * **Keras**: Often used as a higher-level API with TensorFlow, Keras simplifies the creation and training of deep learning models.
2. **Generative Adversarial Network (GAN)**:
   * **Generator**: The SRGAN employs a generator network to create high-resolution images from low-resolution inputs. It is typically a convolutional neural network (CNN) with residual blocks.
   * **Discriminator**: A CNN used to differentiate between real and generated high-resolution images. It guides the generator during training to improve the quality of generated images.
3. **Perceptual Loss Function**:
   * Instead of pixel-wise loss (like Mean Squared Error), SRGAN uses a perceptual loss function based on the VGG network. This helps preserve high-level features and textures in the generated images, resulting in more visually realistic outputs.
4. **Residual Learning**:
   * **Residual Blocks**: SRGAN uses residual learning, which allows for deeper networks without the risk of vanishing gradients, improving the training stability and the final image quality.
5. **High-Performance Computing (HPC)**:
   * **GPUs** (Graphics Processing Units) or **TPUs** (Tensor Processing Units): Due to the high computational cost of training deep learning models, SRGAN models are typically trained on GPUs or TPUs for faster processing.
   * **CUDA**: Utilized for parallel processing during training, helping accelerate the training time significantly.
6. **Image Datasets**:
   * **ImageNet** or **Set5/Set14/B100**: These benchmark datasets are commonly used for training and testing SRGAN models. These datasets contain high-resolution images used to evaluate the performance of image super-resolution models.
7. **Optimization Techniques**:
   * **Adam Optimizer**: A popular optimization algorithm used for training deep learning models. Adam combines the benefits of both Adagrad and RMSProp, providing faster convergence during training.
8. **Data Augmentation**:
   * To improve model generalization, data augmentation techniques like random cropping, flipping, and rotation are often applied to increase the diversity of training data.
9. **Loss Functions**:
   * **Adversarial Loss**: This loss encourages the generator to produce images that the discriminator cannot distinguish from real images.
   * **Content Loss**: Typically a perceptual loss based on features extracted by the VGG network.

**3.DESIGN OF PROPOSED SYSTEM**

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.

**3.1 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]

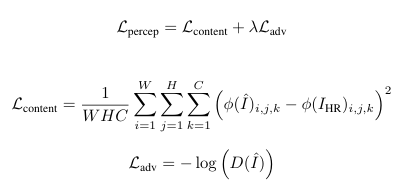
**3.2 Discriminative Evaluation Network**

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]

**3.3 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

**3.4 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.

**3.5 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.

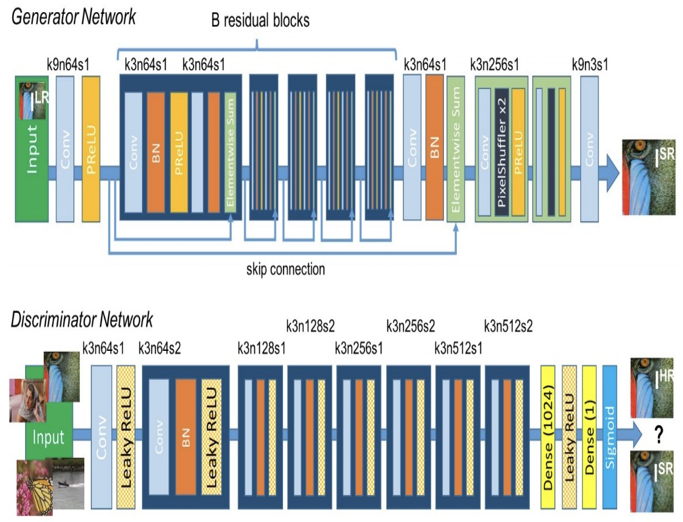


Figure 1 Architecture of SRGAN

**3.6 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.

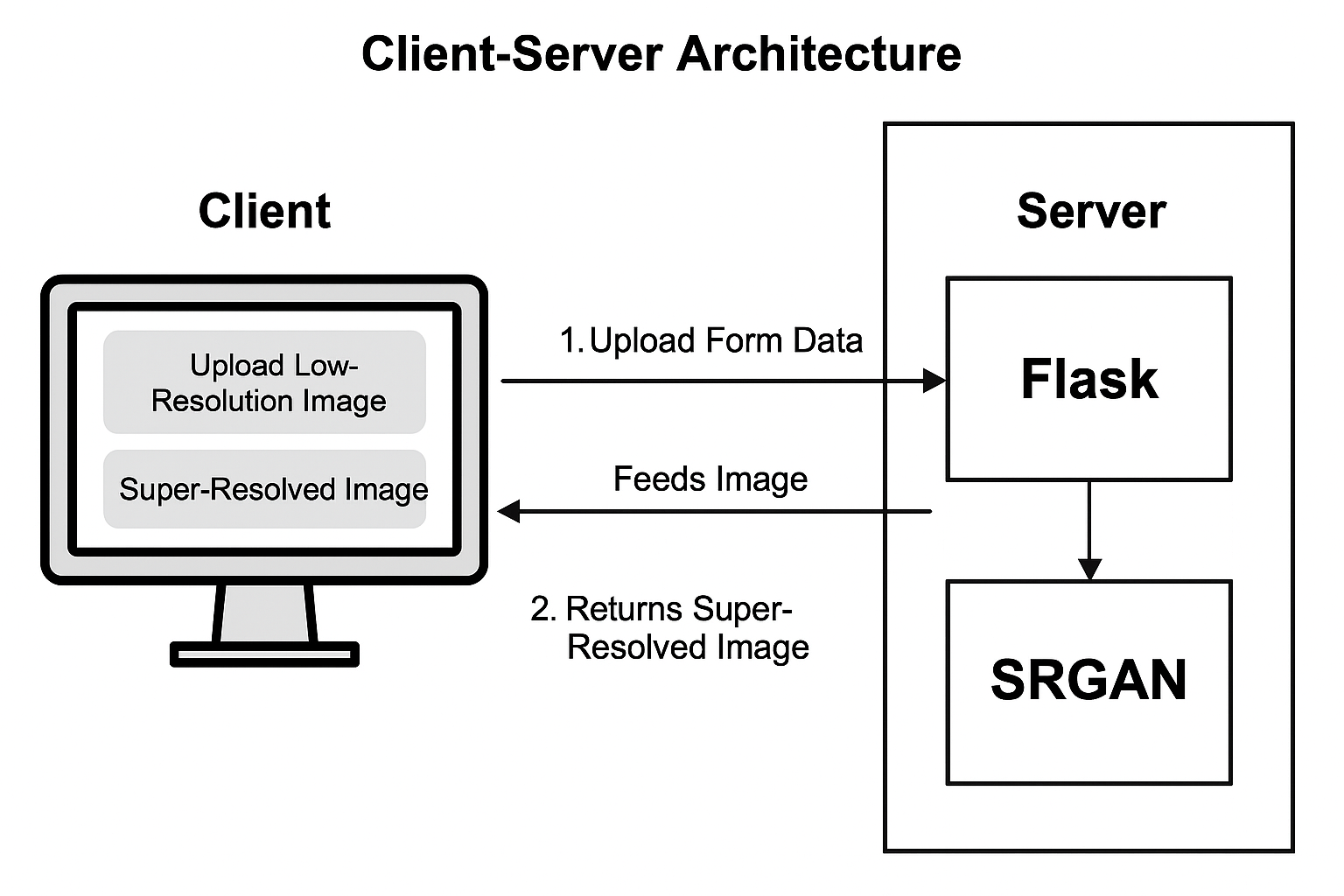
****

Figure 2 Architecture of the Website

**3.7 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.

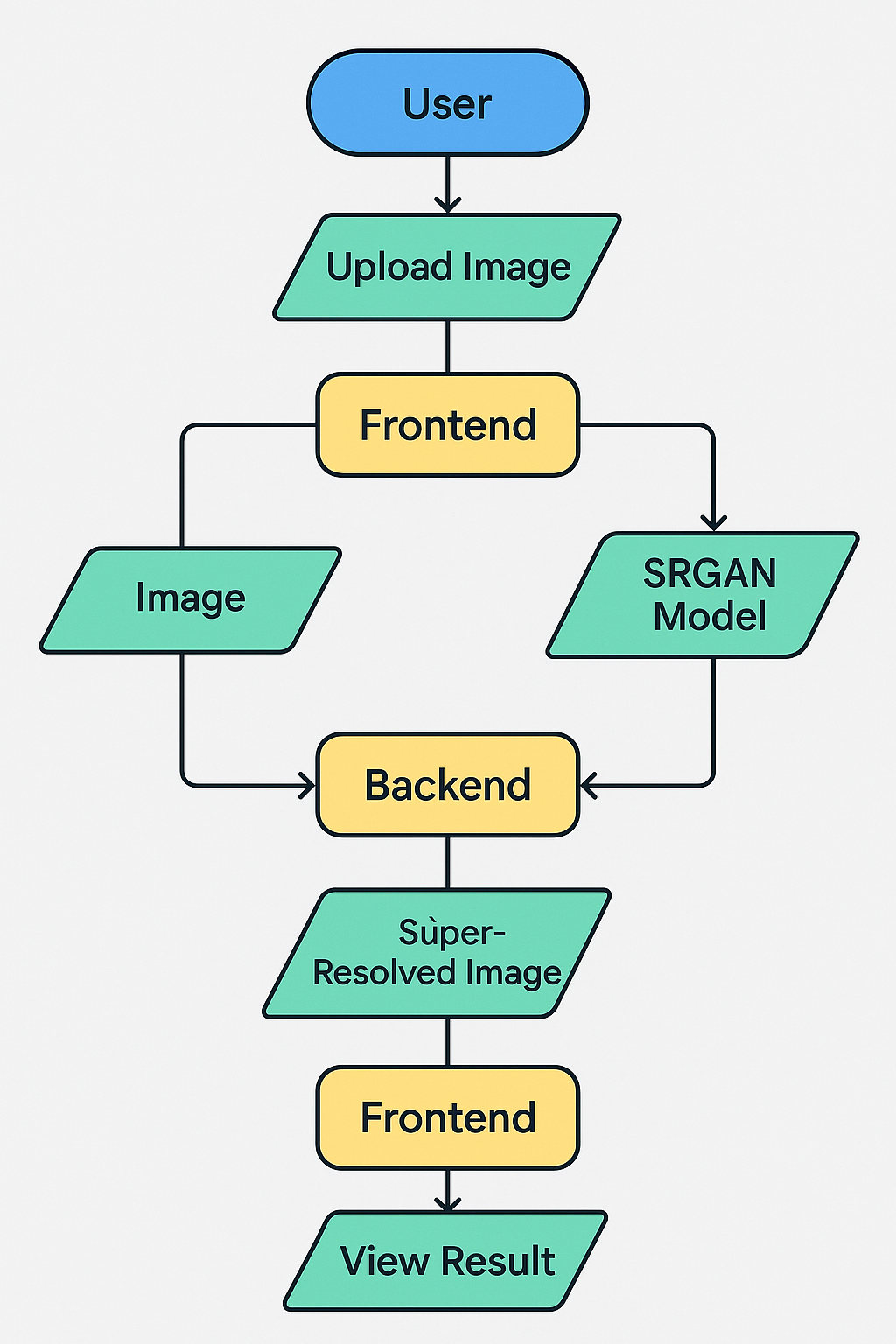


Figure 3 User flow diagram

**4. IMPLEMENTATION OF THE PROPOSED SYSTEM**

**4.1 Module Description**

The proposed system for Image Super-Resolution using SRGAN is divided into the following key modules:

1. **Preprocessing Module:**
   * This module takes the input low-resolution images and processes them before feeding them into the SRGAN model. Preprocessing typically includes normalization (scaling pixel values), resizing images to a consistent size, and converting them into the appropriate format for input into the neural network.
2. **Generator Module**:
   * The core of the system, this module uses a deep residual network (ResNet) to upscale low-resolution images. The generator aims to predict high-resolution images that resemble real images as closely as possible. It uses upsampling techniques, residual blocks, and convolution layers to generate the enhanced image.
3. **Discriminator Module**:
   * This module is designed to differentiate between real high-resolution images and those generated by the generator. It acts as a critic, guiding the generator towards producing more realistic images. The discriminator outputs a probability indicating the "realness" of the generated image.
4. **Loss Calculation Module:**
   * This module computes the loss during training. The loss is a combination of content loss (based on feature maps) and adversarial loss (based on discriminator feedback). The goal is to minimize this combined loss to improve both the generator and discriminator over time.
5. **Training and Optimization Module:**
   * This module handles the iterative training process, using backpropagation and optimization algorithms like Adam to update the parameters of both the generator and discriminator. The weights are updated based on the total loss, including both content and adversarial losses.
6. **Postprocessing Module:**
   * Once the high-resolution image is generated, postprocessing may be applied to refine or enhance the output. This could involve denoising, sharpening, or further image enhancement techniques to improve the visual quality of the generated image.

**4.2 Algorithms**

The primary algorithms used in the implementation of the Image Super-Resolution system using SRGAN are as follows:

1. **Generator Algorithm (ResNet-based):**
   * The generator network is designed to take a low-resolution image as input and output a high-resolution image. It uses residual blocks to capture finer details and upscale the image progressively.
   * Key steps:
     1. Input: Low-resolution image.
     2. Apply residual blocks with skip connections to extract and preserve important features.
     3. Upsample the image using deconvolution layers.
     4. Apply convolution layers to refine image quality.
     5. Output: Super-resolved high-resolution image.
2. **Discriminator Algorithm**:
   * The discriminator is a convolutional neural network (CNN) that distinguishes between real and generated images. It evaluates the generated image and outputs a probability indicating its "realness."
   * Key steps:
     1. Input: Either a real high-resolution image or a generated image.
     2. Apply convolutional layers to extract features from the image.
     3. Output: A probability score (real or fake).
3. **Loss Function Algorithm:**
   * The total loss consists of a combination of content loss and adversarial loss:
     1. Content Loss: Measures the difference between feature maps of the generated image and the real high-resolution image.
     2. Adversarial Loss: Measures how well the generator is at fooling the discriminator.
   * Key steps:
     1. Compute content loss using feature maps from a pre-trained VGG model.
     2. Compute adversarial loss based on the discriminator’s feedback.
     3. Calculate the total loss as the weighted sum of content and adversarial losses.
4. **Optimization Algorithm (Adam):**
   * The Adam optimizer is used to minimize the combined loss during training by updating the weights of the generator and discriminator.
   * Key steps:
     1. Compute gradients of the total loss with respect to the weights of the generator and discriminator.
     2. Update the weights using the Adam optimization algorithm to minimize the loss.

**4.3 Dataset Description**

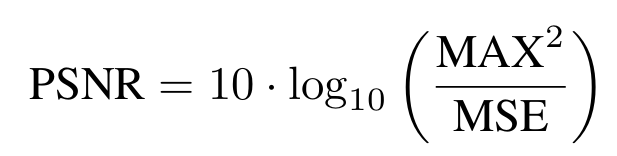
The success of SRGAN depends largely on the dataset used for training. The dataset for this system is a collection of high-resolution images paired with their low-resolution counterparts. These images are often from a variety of domains to make the model generalizable across different use cases.

* **Dataset Characteristics**:
  + **Input Images**: Low-resolution images are generated by downsampling high-resolution images. Typically, a scaling factor (such as 4x downsampling) is used to create low-resolution images from the high-resolution images.
  + **Target Images**: High-resolution images are used as the ground truth for training.
  + **Size**: The dataset should consist of several thousand high-resolution images to ensure the network can generalize well. The more diverse the dataset (e.g., images from various categories like nature, urban, human, etc.), the better the model will perform.
  + **Example Datasets**:
    - **MIRFLICKR Dataset**: Contains high-resolution images for super-resolution tasks and is often used in training SRGANs.
    - **PNEUMONIA XRAY Dataset:** Contains Images of the High-resolution chest X-RAYS
* **Data Augmentation**: Techniques like random cropping, rotation, and flipping are used to augment the dataset and help the model generalize better.

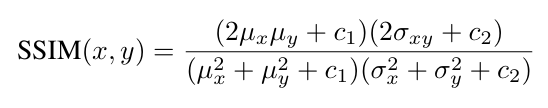
**4.4 TESTING**

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.

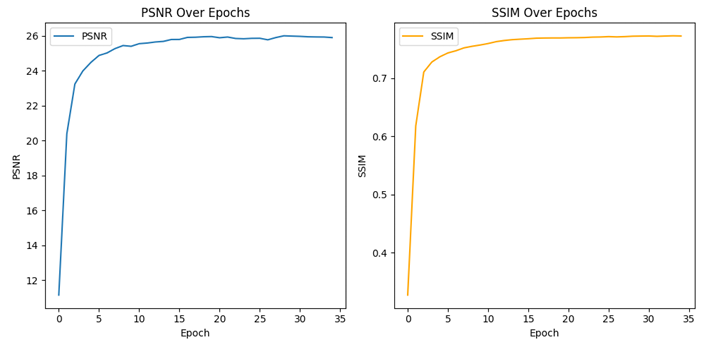
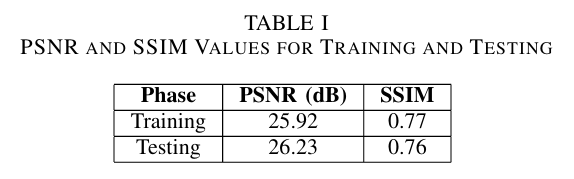


Figure 4 PSNR and SSIM vs Epochs



* **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.

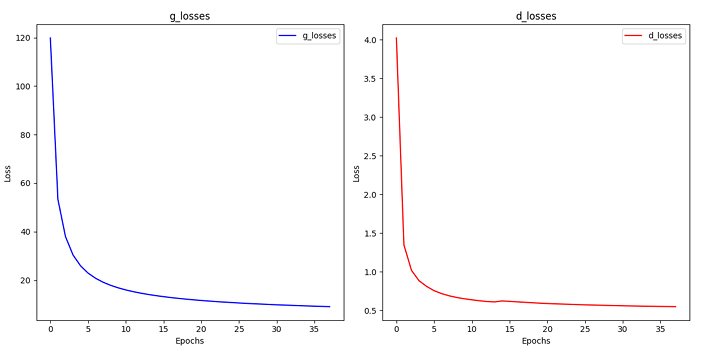


Figure 5 Generator loss and Discriminator loss

**5.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.

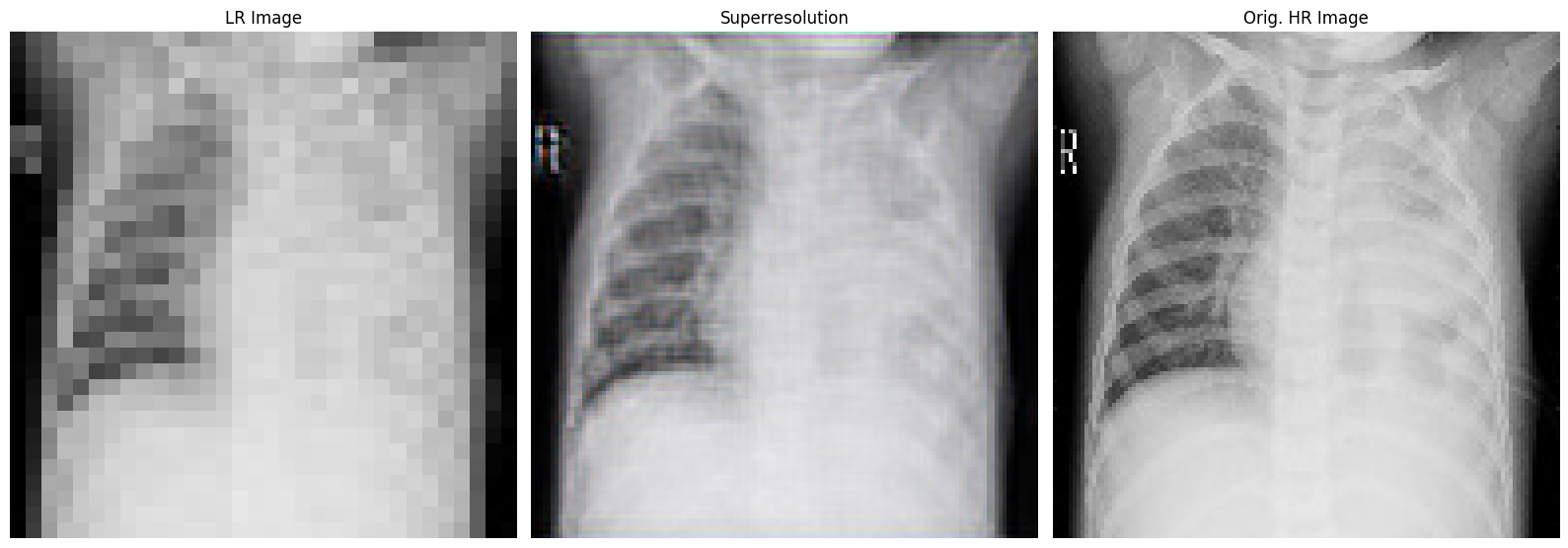
****

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.

**6.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.

**7.REFERENCES**

[1] C. Ledig, L. Theis, F. Huszár, 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 Technology 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 Computer 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. C. 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.

[6] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in Proceedings of the European Conference on Computer Vision (ECCV), 2014.

[7] W. Shi, J. Caballero, M. Theis, F. Huszár, J. Totz, Z. Wang, and W. Freeman, “Real-time single image and video super-resolution using an efficient subpixel convolutional neural network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[8] X. Dong, L. He, H. Zhang, W. Zuo, and M. Zhang, “Image super-resolution with deep convolutional networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[9] M. Aitken, A. Tejani, and W. Shi, “High-quality image super-resolution via deep convolutional networks,” IEEE Transactions on Image Processing, vol. 25, no. 10, pp. 4416–4429, 2016.

[10] T. Lim, S. Son, S. Kim, and S. Kim, “Enhanced deep super-resolution network for single-image super-resolution,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

**8.SUBMISSION OF RESEARCH PAPER**

