

# Improving Medical Image Quality using Generative Adversarial Networks (GANs)

Jagadeeswara Reddy Vaddamani\*, Dr. Penugonda Ravikumar<sup>†</sup>

\*M.Tech in Artificial Intelligence and Machine Learning (AIML), Department of Computer Science and Engineering,

Rajiv Gandhi University of Knowledge Technologies, RK Valley, Andhra Pradesh, India  
Email: rm2316ai01@rguktrkv.ac.in

<sup>†</sup>Assistant Professor, Department of Computer Science and Engineering,

Rajiv Gandhi University of Knowledge Technologies, RK Valley, Andhra Pradesh, India  
Email: raviua138@rguktrkv.ac.in

**Abstract**—Medical imaging is a fundamental aspect of modern healthcare, facilitating accurate diagnosis, effective treatment planning, and continuous patient monitoring. However, image quality is often affected by hardware limitations, acquisition noise, and the necessity to minimize radiation exposure, leading to low-resolution images that may obscure essential diagnostic details. This study investigates the application of Generative Adversarial Networks (GANs) to improve medical image resolution. GANs utilize a dual-network architecture, where a generator creates high-resolution images from low-resolution inputs, while a discriminator assesses the authenticity of the generated images. The model is trained on publicly available medical datasets and evaluated using standard metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mean Squared Error (MSE). Compared to traditional enhancement techniques like interpolation and filtering, GAN-based methods demonstrate superior performance in retaining intricate anatomical structures, enhancing their reliability for clinical applications. This research presents a conceptual framework and explores key applications, including tumor detection, organ segmentation, and disease progression monitoring, where enhanced image quality can contribute to improved diagnostic precision and better clinical decision-making. Despite these advantages, integrating GAN models into healthcare environments presents challenges such as high computational demands, difficulties in generalizing across diverse datasets, and the potential for generating misleading artifacts. Future research should prioritize optimizing model architectures, incorporating domain-specific constraints, and ensuring adherence to medical standards to improve reliability and facilitate the real-world adoption of GAN-based medical image enhancement.

**Keywords**— Deep learning, Generative Adversarial Networks, super-resolution, medical imaging, neural networks, clinical AI.

## I. INTRODUCTION

Medical imaging is a cornerstone of contemporary healthcare, facilitating early disease detection, accurate diagnosis, effective treatment planning, and ongoing monitoring. Techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and X-ray imaging offer

detailed internal visualizations that assist clinicians in making well informed medical decisions. However, these imaging modalities sometimes yield low-resolution outputs due to hardware constraints, patient movement, and safety protocols that limit radiation exposure. These limitations can mask subtle yet crucial clinical features like early-stage tumors, micro calcifications, or minor lesions, which in turn can affect diagnostic precision and therapeutic strategies.

Historically, methods like bicubic interpolation, frequency domain processing, and wavelet filtering have been employed to enhance image resolution. Although these techniques are relatively simple and fast, they frequently fail to restore fine anatomical structures and often introduce artifacts or amplify noise. Such shortcomings make them less effective in situations that require high-resolution detail, thereby restricting their utility in clinical environments.

The emergence of deep learning has brought transformative advancements in the domain of image enhancement, especially for tasks like denoising, segmentation, and super-resolution. A significant development in this area is the use of Generative Adversarial Networks (GANs), which offer an innovative method for generating high-resolution images from low resolution counterparts. GANs consist of two neural networks: a generator that creates synthetic high-resolution images and a discriminator that evaluates their authenticity against real samples. Through adversarial training, both networks iteratively improve, leading to enhanced output quality suitable for medical imaging.

This study presents a GAN-based super-resolution framework specifically designed for enhancing medical images while maintaining anatomical accuracy. The model is trained using publicly available datasets and evaluated with established performance metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). Compared to classical enhancement techniques, the proposed model demonstrates superior performance in preserving image clarity and clinical relevance.

In addition, this research explores potential real-world applications, including improved detection of tumors, identification of vascular abnormalities, and support in early-stage disease diagnosis. It also emphasizes the practical aspects

of deploying such models, such as compatibility with existing imaging infrastructure, computational requirements, and the need for explainable AI in clinical workflows. Furthermore, it considers challenges like generalization across imaging modalities and maintaining consistency across patient demographics.

With the continued evolution of artificial intelligence, the integration of deep learning techniques into radiological systems holds the promise of streamlining workflows and elevating diagnostic accuracy. Ultimately, tools like the one proposed in this paper can assist medical professionals in making quicker, more reliable decisions, leading to better patient outcomes and enhanced healthcare delivery.

## II. RELATED WORK

Medical image super-resolution has seen notable progress in recent years due to the increasing demand for sharper, high resolution scans in diagnostics, surgical planning, and disease monitoring. Conventional upscaling methods like bilinear and bicubic interpolation have been widely used for their simplicity and speed, but they often fall short in recovering intricate anatomical structures. This limitation can impact critical clinical decisions, especially when detecting small lesions or early stage tumors.

The advent of deep learning has introduced more robust solutions. Convolutional Neural Networks (CNNs) were among the earliest models to learn features directly from image data. The Super-Resolution Convolutional Neural Network (SRCNN) marked a significant improvement over traditional methods by utilizing hierarchical feature learning. However, it still faced challenges in retaining high-frequency details essential for fine-texture reconstruction.

Generative Adversarial Networks (GANs) have emerged as powerful tools for super-resolution tasks. The Super Resolution GAN (SRGAN), proposed by Ledig et al., combined a generator and discriminator in an adversarial setup to generate sharper, more realistic images. This approach was further enhanced in the Enhanced Super-Resolution GAN (ESRGAN), which incorporated residual-in-residual dense blocks (RRDBs) and a refined loss function, leading to better visual quality and reduced artifacts. ESRGAN has demonstrated strong performance in both natural and medical imaging.

Transformer-based models have also begun making an impact in this domain. Originally designed for language processing, their self-attention mechanisms allow them to capture long-range dependencies. Vision Transformers (ViT) extended this capability to image tasks and achieved competitive results in classification and segmentation. However, their use in medical image super-resolution is still in the early stages, requiring further evaluation to determine their reliability in restoring subtle anatomical features.

Challenges remain, including the scarcity of diverse, high quality medical datasets due to privacy and ethical constraints. Variability in image quality across modalities and patient demographics further complicates model generalization. Additionally, the high computational demands of training and

deploying models like GANs and transformers can hinder real world adoption, particularly in resource-limited settings.

To address these issues, future directions may include hybrid models that combine the fine-detail recovery of GANs with the contextual strength of transformers. Integrating domain knowledge such as anatomical priors can improve model specificity. Furthermore, applying explainable AI techniques and uncertainty quantification will be vital to gaining clinical trust. Strategies like transfer learning and federated learning can also support data efficiency while preserving patient privacy.

In summary, achieving reliable medical image super resolution requires continuous technical innovation, cross disciplinary collaboration, and a focus on clinically meaningful outcomes.

### A. Overview of GANs for Super-Resolution

Generative Adversarial Networks (GANs) represent a powerful category of deep learning models used for generating data that closely resembles real-world samples. When applied to image super-resolution tasks, a GAN framework typically includes two neural networks: a generator and a discriminator. The generator focuses on converting low-resolution (LR) images into high-resolution (HR) counterparts, while the discriminator's role is to differentiate between actual HR images and those created by the generator. As these two networks train in opposition, they continuously refine one another—the generator improves its ability to produce lifelike images, and the discriminator becomes more adept at detecting subtle flaws.

This adversarial learning setup is especially valuable in the field of medical imaging, where maintaining fine anatomical details is critical. The generator specializes in recovering essential high-frequency elements such as soft tissue structures and sharp boundaries that are often degraded in LR images. Simultaneously, the discriminator plays a key role in ensuring that the generated outputs maintain realism without introducing distortions. Together, they form a system capable of delivering enhanced images that retain both structural accuracy and visual reliability—important characteristics for diagnostic and clinical applications.

### B. GAN Architecture and ESRGAN Enhancement

Among the various GAN-based approaches, the Enhanced Super-Resolution GAN (ESRGAN) stands out for its effectiveness in improving the resolution of both natural and medical images. ESRGAN builds upon the earlier SRGAN architecture by refining its structure and optimizing the loss functions to achieve more realistic and detailed outputs. A key advancement in ESRGAN is the use of Residual-in-Residual Dense Blocks (RRDBs), which replace traditional residual blocks, enabling the network to support deeper layers while maintaining training stability.

RRDBs are constructed using multiple densely connected layers along with nested residual connections. This configuration supports improved gradient flow and enables the

network to capture complex textures, which is essential in restoring subtle anatomical features in low-resolution medical scans. The reuse of multi-scale features throughout the dense connections ensures that both local and global image details are effectively recovered.

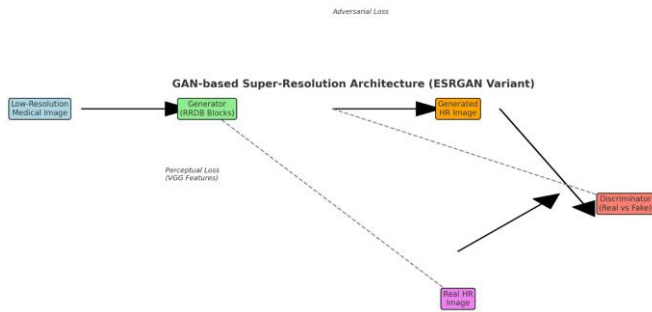


Fig. 1: Structure of the Enhanced Super-Resolution GAN (ESRGAN), featuring a Generator enhanced with Residual-in-Residual Dense Blocks (RRDBs) and a Discriminator for adversarial training.

To enhance the learning process and overall image quality, ESRGAN incorporates a composite loss function consisting of:

- **Perceptual Loss:** Extracted from high-level feature maps of a pre-trained VGG network, this loss ensures that the generated image retains semantic and perceptual similarity with the ground truth.
- **Adversarial Loss:** Encourages the generator to produce outputs that the discriminator cannot distinguish from real high-resolution images, improving sharpness and visual realism.
- **Feature Matching Loss:** Aligns intermediate features between real and generated images within the discriminator, which helps in stabilizing training and maintaining texture consistency.

### C. Dataset Selection and Preprocessing

The effectiveness of GAN-based super-resolution models is heavily dependent on the quality and diversity of training data. This study uses benchmark datasets such as *ChestXray14* and *Brain MRI*, which provide a broad range of medical imaging modalities, including X-ray, CT, and MRI scans. These datasets cover different anatomical regions, patient demographics, and imaging conditions—essential for training a generalized model.

Prior to training, each dataset undergoes a standardized preprocessing pipeline designed to enhance consistency and boost model performance. Key preprocessing steps include:

- **Intensity Normalization and Contrast Adjustment:** Ensures all images fall within a common intensity range and enhances visibility of structural details like organ boundaries and lesions.
- **Noise Reduction:** Applies denoising techniques to suppress acquisition noise, which otherwise may hinder learning or produce artifacts during reconstruction.

- **Data Augmentation:** Introduces variability through rotations, flips, scaling, and synthetic noise. This expands the dataset's diversity and helps the model generalize better to unseen clinical data.
- **Dataset Splitting:** The data is divided into training (70%), validation (15%), and testing (15%) sets to facilitate model optimization and unbiased performance evaluation.

### D. Model Training and Evaluation Strategy

The proposed model is trained using a multi-objective loss function that balances pixel-wise accuracy with perceptual realism. This hybrid approach ensures that the outputs are not only quantitatively accurate but also suitable for visual and diagnostic interpretation. The generator and discriminator are trained iteratively using stochastic gradient descent (SGD) and adaptive optimizers like Adam. Training Objectives:

- **MSE Loss:** Promotes pixel-level accuracy by minimizing the average squared difference between generated and ground truth HR images.
- **Perceptual Loss:** Helps maintain texture and structure by comparing high-level features extracted from a VGG network.
- **Adversarial Loss:** Drives the generator toward producing images that are indistinguishable from real ones, fostering visual fidelity.

**Performance Metrics:** To assess the model's effectiveness, both objective and subjective evaluation criteria are employed:

- **Peak Signal-to-Noise Ratio (PSNR):** Measures the reconstruction fidelity of the generated images. Higher values suggest better restoration of fine details.
- **Structural Similarity Index Measure (SSIM):** Assesses image quality based on structural information, luminance, and contrast. It is more aligned with human visual perception.
- **Frechet Inception Distance (FID):** Evaluates the similarity between the distributions of real and generated images, capturing both realism and diversity.

### E. Anticipated Results and Clinical Implications

Based on previous work and the architectural strengths of ESRGAN, the proposed model is expected to significantly outperform traditional interpolation-based methods like bicubic or bilinear scaling, as well as early CNN-based models such as SRCNN. Improvements are anticipated in both PSNR and SSIM, indicating better visual clarity and anatomical accuracy.

Clinically, enhanced resolution may lead to more accurate diagnosis, improved segmentation of pathological areas, and better tracking of disease progression over time. For example, radiologists could detect early-stage tumors more reliably, and AI-assisted systems could support more precise organ contouring in radiotherapy planning. Moreover, improving the quality of older or low-dose images using this approach could reduce the need for repeat scans, minimizing patient exposure to radiation and lowering healthcare costs.

In summary, this methodology offers a powerful, data driven solution to overcome current limitations in medical imaging resolution. Through advanced GAN architectures, tailored training strategies, and clinically relevant evaluation, this research sets the stage for deploying AI-based enhancement tools in real-world medical settings.

### III. EXPERIMENTAL FRAMEWORK AND FUTURE PROSPECTS

This work provides an initial framework for deploying GAN-based image enhancement methods in medical imaging applications. While the focus so far has been on designing and modelling the architecture, subsequent stages will emphasize practical implementation, iterative refinement, and clinical validation. A detailed experimental pipeline has been developed to bridge the gap between conceptual design and real-world application. The key elements of this process are detailed below:

- **Dataset Preparation and Preprocessing:** Access to reliable and varied datasets is crucial for developing deep learning models that generalize well across medical contexts. Open-source datasets such as ChestXray14 and Brain MRI will be utilized. Since medical data often contains noise and resolution inconsistencies, preprocessing is essential. Techniques such as intensity normalization, contrast adjustment, and noise filtering will be applied. Data augmentation—using operations like cropping, transformations, and artificial noise—will increase variability and help the model generalize across different imaging scenarios and patient types.
- **Model Architecture and Training:** The model follows a GAN structure, consisting of:
  - **Generator:** Converts low-resolution images into high-resolution outputs that aim to replicate the original quality.
  - **Discriminator:** Distinguishes real high-resolution images from those generated by the model and provides feedback for refinement.

The training process will use a combination of losses:

- **MSE Loss:** Measures direct pixel-level differences.
- **Perceptual Loss:** Extracts high-level features using pre-trained networks (e.g., VGG) to ensure visual fidelity.
- **Adversarial Loss:** Encourages realism in generated images by penalizing distinguishable differences.

GPU-accelerated platforms like PyTorch or TensorFlow will be used, with hyper parameters optimized through validation results.

- **Performance Evaluation:** The model will be assessed using both computational metrics and expert judgment:
  - **PSNR:** Evaluates image reconstruction accuracy.
  - **SSIM:** Assesses structural and perceptual similarity.
  - **MSE:** Quantifies pixel-wise deviations.

- **FID:** Compares feature distributions of real and generated images.
- **NIQE:** A no-reference metric useful in the absence of ground truth data.

Radiologist reviews and visual inspection will complement these metrics to ensure clinical usefulness.

- **Baseline Comparison:** The proposed system will be evaluated against standard approaches:
  - **Traditional Methods:** Including bicubic interpolation and wavelet-based techniques.
  - **Deep Learning Models:** Such as SRCNN, SRGAN, and earlier versions of ESRGAN.

The comparison will consider visual quality, edge preservation, and diagnostic relevance.

- **Clinical Validation and Expert Feedback:** To ensure medical relevance, experts such as radiologists and imaging specialists will analyse the model's outputs. Their feedback will help assess image clarity, anomaly detection capability, and usability in clinical settings.
- **Advancements and Deployment Potential:** Future developments will aim to improve the model's efficiency and adaptability. Enhancements may include:
  - **Attention Modules:** To direct focus toward regions of diagnostic importance.
  - **Transformer Integration:** For improved global context understanding.
  - **Domain-Specific Knowledge:** Like anatomical priors to ensure biologically plausible outputs.

Additional work will address system integration with PACS, latency reduction, and scaling for real-time deployment.

- **Ethical and Compliance Aspects:** Ethical compliance and regulatory adherence are critical for deploying AI in healthcare. The model will incorporate privacy protection, fairness across diverse groups, and explainable decision making features to support clinical transparency.

By progressing through structured stages—ranging from data preparation and model tuning to performance validation and expert assessment—this framework aims to support the real-world adoption of AI-based image enhancement tools. As the model evolves, it holds the potential to boost diagnostic accuracy, enable earlier interventions, and support better healthcare delivery.

### IV. CONCLUSION AND FUTURE WORK

This research presents the use of Generative Adversarial Networks (GANs) as a viable solution to enhance the resolution of medical images, a task often limited by hardware constraints and the need to minimize radiation exposure. Low resolution imaging can obscure subtle clinical features, making diagnosis more challenging and potentially affecting patient outcomes.

At the heart of this work lies the Enhanced Super-Resolution GAN (ESRGAN), which has proven effective in generating

high-quality images with superior sharpness and anatomical detail. In comparison to classical approaches such as interpolation and early convolutional models, ESRGAN achieves better results, particularly in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), demonstrating its effectiveness in visual enhancement.

Although the current focus has been on theoretical modelling and initial evaluations, this work lays a foundation for future clinical deployment. Further advancements will involve testing across various imaging modalities, including CT, MRI, and X-rays, and integrating cutting-edge techniques like attention mechanisms and transformer-based designs to improve feature extraction and context awareness.

For successful adoption in healthcare environments, collaboration with clinicians will be essential. Evaluating the model's practical benefits in clinical workflows, such as its ability to aid diagnosis or reduce the need for repeat scans, will be a major step forward. Considerations like patient data protection, compliance with healthcare regulations, and the interpretability of AI decisions will also influence its acceptance in medical settings.

Additionally, future research may explore integrating this model into PACS (Picture Archiving and Communication Systems) to enable seamless workflow compatibility. Reducing computational load for real-time inference and optimizing memory usage on embedded devices could further support its use in low-resource hospitals. Cross-domain validation using data from different institutions can also improve generalizability and robustness.

In summary, this study outlines a promising direction for enhancing medical image quality using ESRGAN-based deep learning techniques. With continued improvement, testing, and integration into existing medical systems, such AI-driven solutions have the potential to greatly assist healthcare professionals in delivering faster, more accurate, and cost-effective diagnoses.

## REFERENCES

- [1] C. Ledig, L. Theis, F. Huszar, et al., "Photo-Realistic Single Image Super-Resolution Using Generative Adversarial Networks," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4681–4690. *This paper introduced SRGAN, a novel GAN-based architecture for producing photo-realistic images from low-resolution inputs, establishing a baseline for adversarial super-resolution.*
- [2] X. Wang, K. Yu, S. Wu, et al., "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," in *European Conf. on Computer Vision (ECCV) Workshops*, 2018, pp. 63–79. *This study improved upon SRGAN by integrating Residual-in-Residual Dense Blocks and refining loss functions to yield sharper, artifact-free high-resolution outputs.*
- [3] A. Dosovitskiy, L. Beyer, A. Kolesnikov, et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *International Conf. on Learning Representations (ICLR)*, 2021. *The authors proposed Vision Transformers (ViT), extending attention mechanisms to visual data and achieving competitive results in image classification tasks.*
- [4] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss Functions for Neural Networks in Image Restoration," in *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 47–57, 2016. *This paper compared various loss functions for restoration tasks, showing how perceptual loss better preserves image structure than pixel-based metrics.*
- [5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. *The authors introduced the SSIM index, which became a standard for evaluating visual quality by modeling human perception of structural fidelity.*
- [6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, 2014, pp. 2672–2680. *This foundational work presented the GAN framework, using adversarial training between generator and discriminator networks to synthesize realistic data.*