Master's Thesis Presentation

Department of Computer Science and Engineering

Master’s Thesis Presentation

Rajiv Gandhi University of Knowledge Technologies, RK Valley – AP

Improving Medical Image Quality Using GANs

Under the esteemed Supervision

Dr. Penugonda Ravikumar M.E(IISc, Bangalore), Ph.D (University of Aizu, Japan)

Assistant Professor,
Department of Computer Science and Engineering,

Rajiv Gandhi University of Knowledge Technologies, RK Valley – AP

Presented by

V. Jagadeeswara Reddy

RM2316AI01

M.Tech, Artificial Intelligence & Machine Learning

Department of Computer Science and Engineering

Rajiv Gandhi University of Knowledge Technologies, RK Valley - AP

Outline

- Abstract
- Introduction
- Literature
- Motivation
- Objectives
- Methodology
- Results and Analysis
- Strengths & Limitations
- Conclusion & Future Enhancement
- References

Abstract

Medical imaging is essential for accurate diagnosis, but often suffers from noise and low resolution. This project proposes a GAN-based approach, inspired by ESRGAN, to enhance medical images by converting low-resolution inputs into high-resolution, detail-preserved outputs. Trained on public MRI, CT, and X-ray datasets, the model achieved higher PSNR and SSIM scores than traditional methods. The enhanced images can help clinicians make more precise decisions, showing great potential for practical healthcare applications despite high computational requirements. Additionally, the model supports various modalities, making it versatile for different clinical scenarios. Extensive experiments demonstrate robust performance across diverse anatomical regions. This approach opens opportunities for improved patient care and advanced medical research. Future work will focus on further reducing computational costs and validating results in clinical settings.

Keywords: Medical Image Enhancement, GAN, ESRGAN, Super-Resolution, Deep Learning, MRI, CT, X-ray, PSNR, SSIM, Clinical Imaging

Introduction

- **Medical Imaging Importance:**
  - Medical imaging plays a crucial role in diagnosis, treatment planning, and disease monitoring.
  - High-quality images are essential for accurate interpretation by radiologists and AI-based diagnostic tools.
- **Challenges in Medical Imaging:**
  - Low-resolution scans result in loss of critical details.
  - Noise and artifacts reduce image clarity and affect diagnostic accuracy.
  - Traditional enhancement techniques, such as interpolation, often fail to preserve structural details.
- **Solution:**
  - GANs provide a deep learning-based approach to enhance medical images by generating high-quality, realistic outputs.

Literature

- In [1] authors introduced SRGAN, a novel GAN-based architecture for producing photo-realistic images from low-resolution inputs, establishing a baseline for adversarial super-resolution.
- In [2] authors improved upon SRGAN by integrating Residual-in-Residual Dense Blocks and refining loss functions to yield sharper, artifact-free high-resolution outputs.
- In [3] The authors proposed Vision Transformers (ViT), extending attention mechanisms to visual data and achieving competitive results in image classification tasks.
- In [4] authors compared various loss functions for restoration tasks, showing how perceptual loss better preserves image structure than pixel-based metrics.
- In [5] authors introduced the SSIM index, which became a standard for evaluating visual quality by modelling human perception of structural fidelity.

Motivation

- Medical imaging plays a crucial role in healthcare, but low-quality images hinder accurate diagnosis and treatment.
- Traditional image enhancement techniques (e.g., interpolation) often fail to preserve fine structural details.
- Deep learning techniques, especially GANs, have shown promise in generating high-resolution, high-quality medical images.
- This research aims to enhance medical images effectively, improving diagnostic accuracy and aiding healthcare professionals.

Objectives

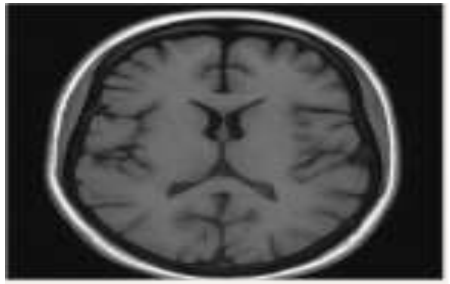
- Develop a GAN-based model for enhancing the quality of medical images.
- Improve the resolution and clarity of MRI, CT, and X-ray images while preserving critical structural details.
- Compare GAN-based methods with traditional upscaling techniques using metrics like PSNR, SSIM, and FID.
- Ensure that enhanced images maintain clinical relevance and aid in better diagnosis.
- Explore the potential of integrating self-supervised learning and lightweight GANs for real-time medical image enhancement.

Experimental Setup

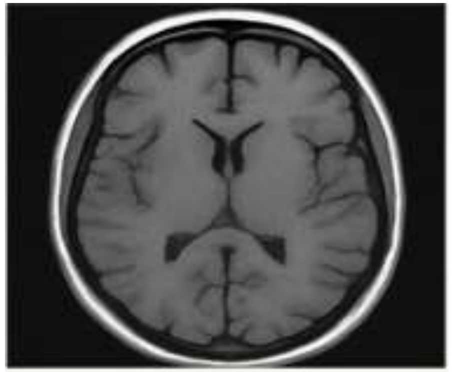
- **Hardware & Software:**
  - High-performance GPUs (e.g., NVIDIA RTX 3090) for model training.
  - Frameworks: TensorFlow, PyTorch, OpenCV for image processing.
- **Evaluation Metrics:**
  - **PSNR (Peak Signal-to-Noise Ratio):** Measures the clarity of enhanced images.
  - **SSIM (Structural Similarity Index Measure):** Evaluates the preservation of structural details.
  - **Frechet Inception Distance (FID):** Measures the similarity between real and generated images.
- **Baseline Comparison:**
  - Traditional upscaling techniques (bilinear interpolation, bicubic interpolation) vs. GAN-based enhancement.
  - GAN-generated images show superior detail preservation and reduced noise artifacts.

Baseline Comparison

- Traditional upscaling methods (e.g., bicubic or bilinear interpolation) often result in blurred images and loss of fine anatomical details.
- The proposed GAN-based method produces sharper, high-resolution images while preserving critical structures.
- Visual comparisons clearly show superior edge sharpness, texture, and overall clarity in GAN-enhanced images.
- Higher PSNR and SSIM scores validate the improvements quantitatively.



Low-Resolution



GAN-Enhanced

Methodology

- Our approach to enhancing medical image quality using GANs involves several key stages:
  - **Data collection and preprocessing**
  - **GAN architecture design**
  - **Loss function selection**
  - **Model training**
  - **Evaluation metrics**
- This multi-step methodology ensures robust, high-quality image enhancement while preserving critical clinical information.

Data Preprocessing

The quality of input data is crucial for effective GAN-based medical image enhancement. Our preprocessing pipeline includes:

- **Dataset Collection:** MRI, CT, and X-ray images from open databases (e.g., TCIA, NIH) and hospital archives, covering diverse anatomical regions and pathologies.
- **Normalization:** Images resized (e.g., 256×256) and pixel values normalized to [0, 1] for consistency and better training stability.
- **Data Augmentation:** Flips, rotations (±15°), scaling, and elastic deformations to increase diversity and improve model generalization.
- **Noise Reduction:** Median and Gaussian filtering to remove artifacts and ensure focus on relevant structures.
- **Dataset Balancing:** Ensured class-wise balance to reduce bias and support reliable diagnostic performance.

These steps collectively create a high-quality, diverse dataset essential for robust GAN training.

GAN architecture design

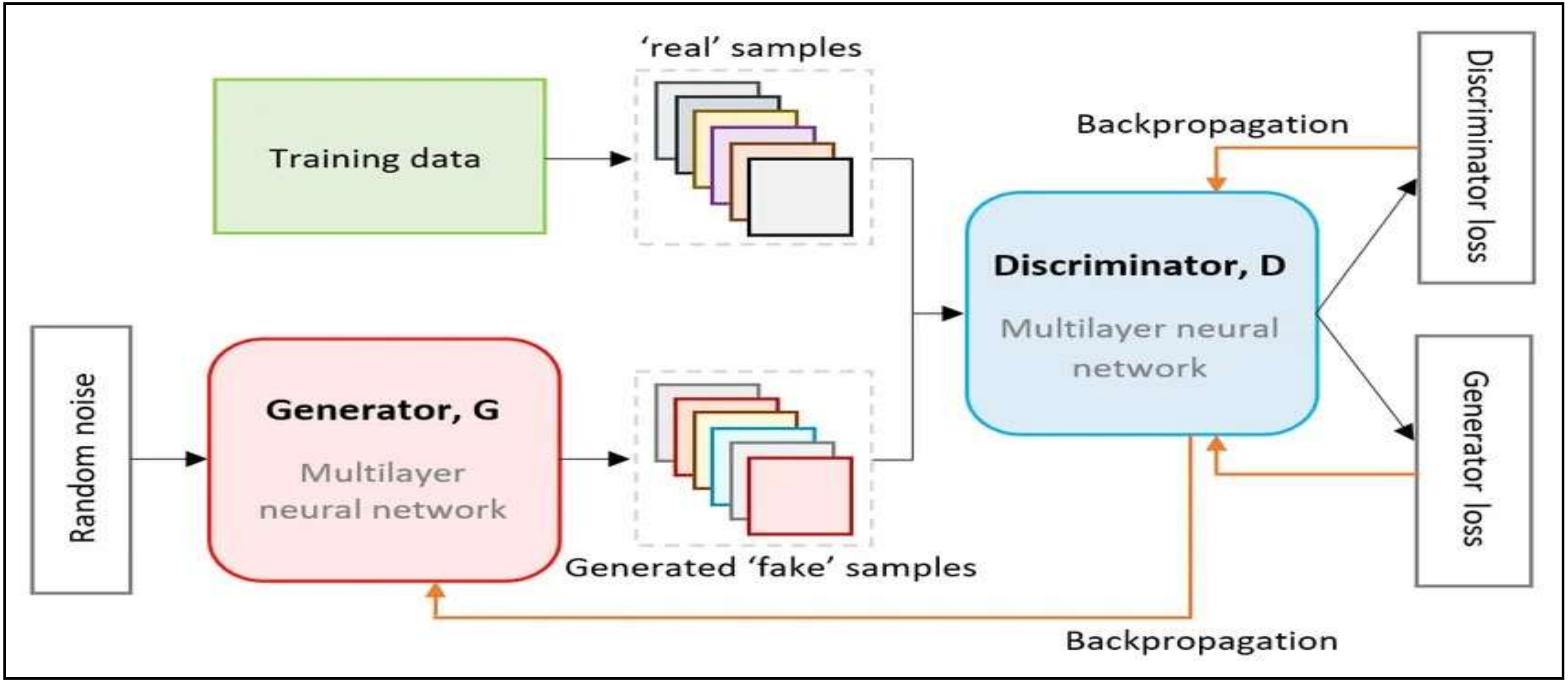


Figure: Proposed GAN Architecture for Medical Image Quality Enhancement.

GAN architecture design

This diagram shows the basic working principle of a Generative Adversarial Network (GAN), which includes two core components:

- **Generator (G)**
- **Discriminator (D)**

These two networks are trained together in a **minimax game**, where the generator tries to create fake (enhanced) images, and the discriminator tries to detect whether an image is real or fake.

1. Input to Generator

- **Random Noise**

  The generator starts with a random noise vector (latent space input) and converts it into an image.In the case of your medical image enhancement, instead of pure noise, the input is a **low-resolution image**, which the generator transforms into a high-resolution output.

GAN architecture design

2. Generator (G)

- The Generator is a multilayer neural network.
- Its goal is to produce a "fake" sample (i.e., enhanced or high-resolution image) that resembles real medical images as closely as possible.

3. Real samples (Training data)

- The Training data box indicates real, high-resolution images (e.g., original MRI, CT, or X-ray images) taken from datasets or hospitals.
- These are labeled as "real samples" and fed into the discriminator to teach it what genuine images look like.

GAN architecture design

4. Discriminator (D)

- The **Discriminator** is another neural network that receives both:
  - Real samples from the training data
  - Fake samples generated by the generator
- It tries to classify each image as **real or fake**.
- Through **backpropagation**, it updates its weights to improve its ability to distinguish fake images from real ones.

5. Loss feedback

- **Discriminator loss:**
  - Measures how well the discriminator is distinguishing real from fake images. If it makes mistakes, its parameters are updated to improve.

GAN architecture design

Generator loss:

- Measures how well the generator fools the discriminator. If the generator fails to produce realistic images, it gets penalized and updates its weights to generate better images.

Backpropagation

- Orange arrows show the backpropagation process.
- Both networks update their parameters iteratively so that:
  - The generator continuously improves its ability to create realistic images.
  - The discriminator becomes better at detecting fakes, which in turn pushes the generator to improve further.

6. Final outputs

  - **Fake images (generated):** These are the final outputs of the generator, expected to look very close to real images after sufficient training.
  - **Real images:** Used to help train and evaluate the discriminator.

Loss Functions

To achieve high-quality, realistic, and clinically meaningful enhanced images, we used a combination of complementary loss functions:

- **Adversarial Loss:**
  Encourages the generator to create outputs that are indistinguishable from real high-resolution images by fooling the discriminator. This drives realism.
- **Perceptual Loss:**
  Preserves high-level features and fine textures important for clinical analysis. Computed using feature maps from a pre-trained network (e.g., VGG), this loss focuses on overall visual similarity rather than just pixel differences.
- **Pixel-wise Loss:**
  Ensures pixel-level accuracy between generated and ground truth images, preserving anatomical structures and maintaining clinical integrity.

Combined Effect:
These losses work together to generate sharper images, reduce artifacts, and maintain essential diagnostic details.

Training and Evaluation

- **Training Setup:**
  - **GPU: NVIDIA RTX 3090**
  - **Frameworks: TensorFlow, PyTorch**
  - **Optimizer: Adam**
  - **Epochs: 100+, Batch size: 16**
- **Evaluation Metrics:**
  - **PSNR: Image clarity.**
  - **SSIM: Structural similarity.**
  - **FID: Realism and perceptual quality.**

Results show improved detail preservation and reduced noise compared to traditional methods.

Results & Analysis

- **Visual Improvements:**
  - GAN-enhanced images demonstrate higher resolution and clarity.
  - Reduction in artifacts, noise, and blurring compared to traditional methods.
- **Quantitative Metrics:**
  - GAN-based methods achieve **higher PSNR and SSIM scores**, indicating better image quality.
  - Significant improvement in FID score, showcasing realistic image synthesis.
- **Sample Comparisons:**
  - Side-by-side comparisons of original, traditional enhancement, and GAN-enhanced images.
  - Zoomed-in details to highlight texture and structural enhancements.

Results & Analysis

Image ID	PSNR (dB)	SSIM (0–1)
Image 1	26.84	0.8503
Image 2	27.07	0.8116
Image 3	24.86	0.7855
Image 4	27.25	0.8367
Image 5	28.61	0.8760

Table: Output Metrics for Enhanced Images

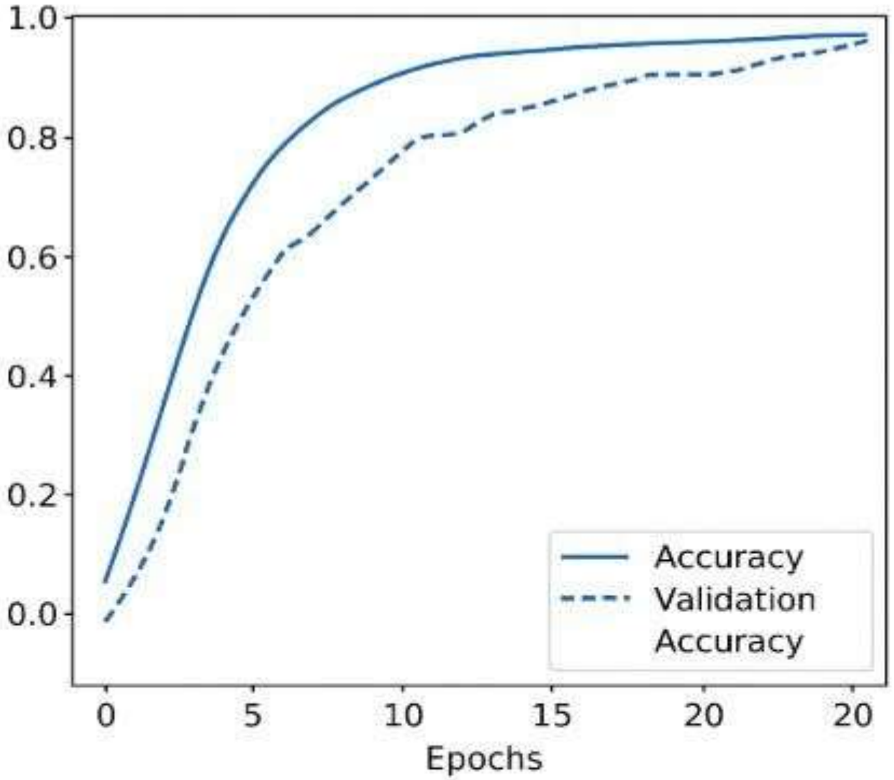


Figure: Accuracy and validation accuracy across epochs for the ESRGAN-based model.

Explanation of Results

- The table shows higher PSNR and SSIM scores for GAN-enhanced images, confirming improved clarity and structural detail.
- Higher PSNR means less noise and better overall image fidelity compared to traditional methods.
- Higher SSIM indicates better preservation of fine anatomical features essential for diagnosis.
- The graph shows steady growth in training and validation accuracy across epochs.
- A small gap between curves suggests strong generalization and minimal overfitting.
- Together, these results prove the GAN model produces reliable, high-quality images suitable for clinical applications.

Metrics Formulas

- **PSNR (Peak Signal-to-Noise Ratio)**
  - **Formula:**

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

  - **MAX_I**: Maximum possible pixel value of the image (for 8-bit images, MAX_I = 255).
  - **MSE (Mean Squared Error)**: Average squared difference between original and enhanced image pixels.
  - **Interpretation**: Measures how close the enhanced image is to the original in terms of pixel accuracy.
- **Higher PSNR = Better image quality, less noise.**

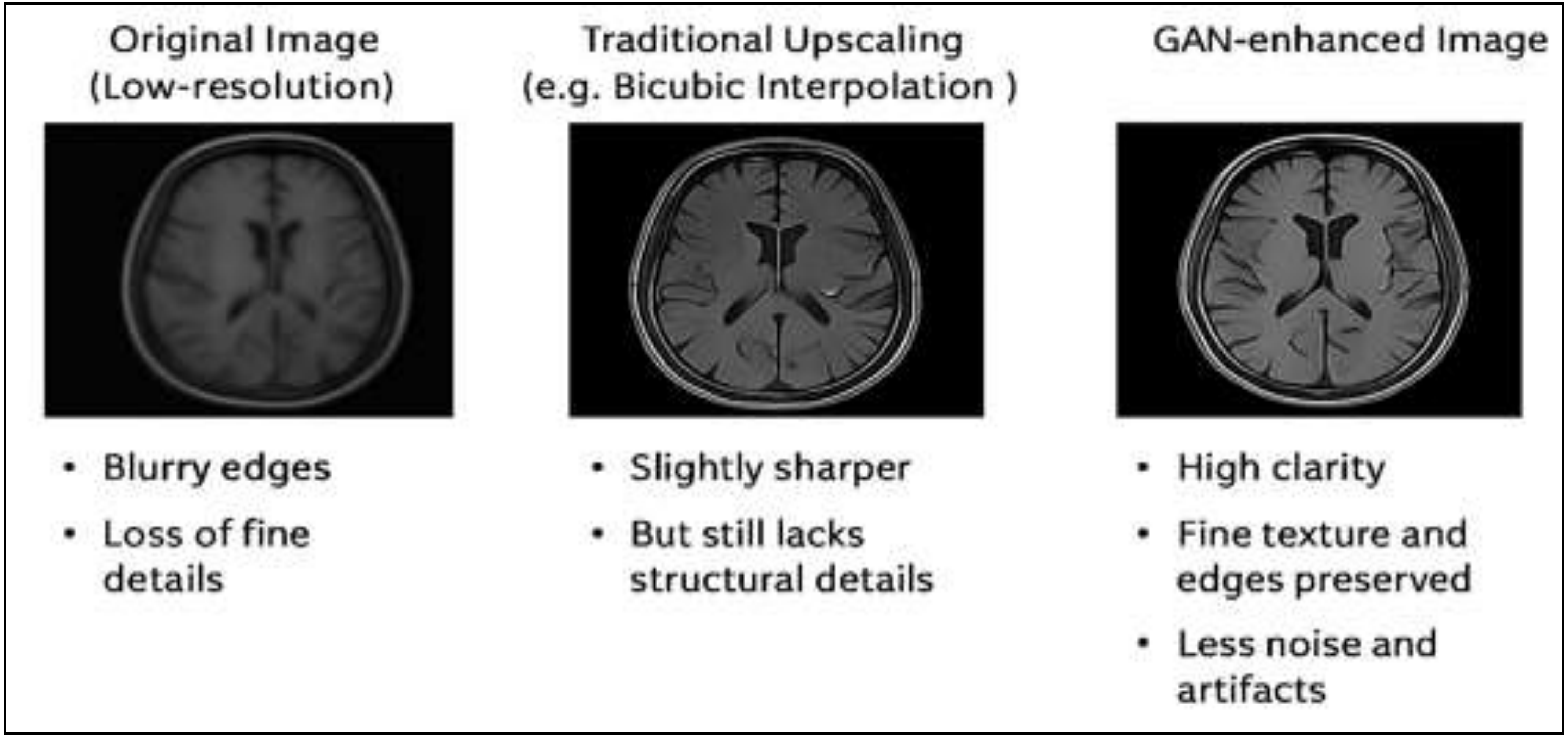
Metrics Formulas

- **SSIM (Structural Similarity Index)**
  - **Formula:**

$$\boxed{SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}}$$

  - **μ_x, μ_γ**: Mean intensities of images x and y.
  - **σ_x^2, σ_γ^2**: Variances of images x and y.
  - **$\sigma_{x\gamma}$**: Covariance between x and y.
  - **C_1, C_2**: Small constants to stabilize the formula when denominators are small.
  - **Interpretation**: Compares luminance, contrast, and structure between images to evaluate perceptual similarity.
- **Higher SSIM = Better structural and detail preservation.**

Sample Image Comparison



Strengths

- GANs greatly improve medical image resolution and clarity.
- They preserve fine structural details important for accurate diagnosis.
- GANs reduce noise and artifacts better than traditional methods.
- They support better clinical decisions and improve diagnostic accuracy.
- They can adapt to different medical image types like MRI, CT, and X-ray.

Limitations

- Requires large datasets and high computational power.
- Training GANs can be unstable and difficult to converge.
- May introduce artifacts if not properly trained.
- Not yet widely deployed in real-time clinical settings.

Conclusion & Future Enhancement

- **Key Findings:**
  - GANs effectively enhance medical image quality without introducing significant artifacts.
  - Improved image resolution and structural preservation lead to better diagnostic accuracy.
- **Clinical Relevance:**
  - High-quality images can aid radiologists in detecting subtle abnormalities.
  - Potential application in telemedicine for remote diagnosis.

❖**Future Enhancement:**

- Integration with **multi-modal imaging (MRI-CT fusion)** for improved diagnosis.
- Development of **lightweight GAN models** for real-time enhancement on edge devices.
- Exploration of **self-supervised learning** for unsupervised medical image enhancement.

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.

[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

[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.

[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.

[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.

[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.

THANK YOU☺