Enhancing Medical Image Quality Using GANs with Evolutionary Hyperparameter Optimization

AUTHORS:
SWARA JOSHI, TEERTH CHAUHAN & SIA NAIR

DEPARTMENT OF ARTIFICIAL INTELLIGENCE
MPSTME, NMIMS UNIVERSITY, MUMBAI

UNDER THE GUIDANCE OF:
DR. AMI MUNSHI

***Abstract* —** Image enhancement is a vital step of medical image analysis and image recognition. X-ray and ultrasound imaging are the most preferred medical imaging technologies which are important for diagnosis of disease. But the edges and borders on image are not as clear as expected due to interference and low intensity in images. This paper presents an images enhancement techniques, specially in the case of medical images. Using image enhancement, it is possible to get the details which are kept hidden as well as to improve the image contrast. In the case of analyzing image, the commencing part is that the edge of an image. Successful results of image analysis depend on edge detection & enhancement. In this research we have developed a method of enhancement by incorporating a U-Net-based Generative Adversarial Network (GAN). Additionally, we integrate Evolutionary Strategy (ES) to optimize critical hyperparameters, enhancing model performance significantly. This developed method is tested on low contrast medical images and by observing the results it can be said that this applied methods perform well for enhancing medical images.

I. INTRODUCTION

The quality of medical images is paramount for accurate diagnostics and treatment planning. However, real-world medical images — particularly MRI brain scans — are often affected by low contrast, noise, and resolution loss due to acquisition limitations and environmental factors. Enhancing such images is not merely about aesthetic improvement but also about preserving the structural and diagnostic integrity critical for clinical decisions.

Several traditional image enhancement techniques exist, such as histogram equalization, adaptive filtering, and contrast stretching. While effective in improving global contrast, they often fail to preserve local structural details or introduce artifacts. Recently, deep learning-based techniques, especially Generative Adversarial Networks (GANs), have shown promise in medical imaging tasks, including super-resolution, denoising, and enhancement.

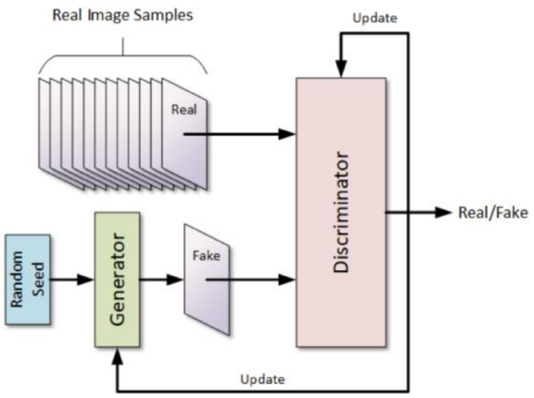
In this study, we employ a U-Net-based Generator and PatchGAN Discriminator under the Pix2Pix framework. This setup allows for better feature preservation and spatial accuracy. Additionally, we introduce an Evolutionary Strategy to tune critical generator hyperparameters like filter size, kernel dimensions, and dropout probability, aiming to balance perceptual realism with structural fidelity.

Our proposed system undergoes multiple evaluations including visual comparison and quantitative assessments using PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). The results demonstrate notable improvement in image clarity and diagnostic value after enhancement. The goal is to present a pipeline that can assist radiologists and machine learning systems in analyzing brain MRI data more effectively.

II. METHODS

II.1 GAN ARCHITECTURE

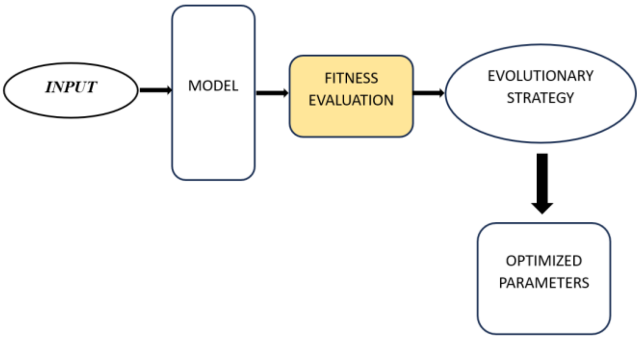
Generative Adversarial Networks, pioneered in 2014 by Goodfellow et al.[1], introduced a sophisticated architecture that went beyond the available capabilities of contemporary generative models. The GAN framework hinges on the contest between two distinct neural networks: the generator G and the discriminator D. Unlike other non-generative models, GANs are not based on minimizing the loss on a single set of input data. Operating in game like environment, these networks G and D have opposing objectives. GANs are formulated as a zero-sum minimax game, where the loss functions of each player is balanced by the loss of other player, and the system tries to reach Nash equilibrium between the loss functions. Generative Adversarial Network Architecture The “Generator”, as the name suggests, is dedicated to creating synthetic data that mirror the characteristics of real datasets. The role of the “Discriminator” is to try and discern between authentic and generated (fake) data, by evaluating the accuracy of the generator's outputs. As the training progresses, the generator iteratively refines its outputs in a bid to deceive the discriminator. This adversarial dynamic results in the generator producing high-fidelity data, almost indistinguishable from real data. The convergence of this iterative process to a high degree is key to the GANs success as compared to other models. Unlike other models, GANs precision does not reduce when a small amount of noise is added to the original data (which can happen due to overfitting or insufficient training dataset). GANs do still suffer from instability issues like mode-collapse or non-converging gradients. This is an area of ongoing research.



Generative Adversarial Networks (GANs) have revolutionized image recognition by enabling high-resolution synthetic image generation and the ability to perform style transfers. GANs can skillfully perform advanced tasks like super-resolution with a very high degree of precision. Super resolution refers to the technique of enhancing the resolution of an image beyond its original dimensions, thereby producing a higher definition and clearer representation of the original image. Often real-world images lack sharp details and super-fine textures. This process not only increases the pixel density but also preserves the intricate details, textures, and patterns within the original image.

II.3 Evolutionary Strategy for Hyperparameter Tuning

Evolutionary Strategies (ES) are a subset of bio-inspired optimization techniques derived from the principles of natural selection and evolution. These strategies are particularly effective for solving complex, non-convex optimization problems, such as tuning deep learning models with high-dimensional search spaces.
In this project, we leveraged ES to automatically optimize key generator hyperparameters that significantly impact the model's ability to enhance medical images. The goal was to improve training stability and output quality without relying on manual grid or random search, which are often inefficient and computationally expensive.



The hyperparameters optimized include:

- **Number of convolutional filters**: Controls the feature extraction capacity.
- **Kernel size**: Influences the receptive field and local spatial learning.
- **Dropout rate**: Helps prevent overfitting during training.

The ES process started with a randomly initialized population, where each individual encoded a unique set of the above parameters. A fitness function was defined using the generator loss from a short training cycle (typically 1–3 epochs). Each generation underwent mutation (random adjustments to parameters), and selection favored individuals with lower generator losses.

II.3 CLAHE

Contrast Limited Adaptive Histogram Equalization (CLAHE). CLAHE is an advanced contrast enhancement technique that operates on small image regions (tiles) rather than the entire image, thereby improving local contrast while minimizing the amplification of noise. CLAHE limits contrast amplification in homogeneous areas by applying a clip limit to the histogram before redistributing it.

The mathematical formulation of CLAHE involves:

- Dividing the image into contextual regions.
- Computing the histogram of each region and clipping it using a predefined threshold.
- Redistributing the clipped pixels uniformly across all histogram bins.
- Applying bilinear interpolation to merge neighbouring tiles smoothly and avoid artificial boundaries.

Let H(i) be the histogram value of intensity level i, and T be the clip limit. The clipped histogram Hc(i) is defined as:

- Hc(i) = min(H(i), T) ---- 1

The excess pixels E = Σ(H(i) - T)+ are then redistributed equally:

- H_redistributed(i) = Hc(i) + E / N ---- 2

III. PROPOSED APPROACH

The proposed framework integrates several stages ranging from data acquisition to deep learning-based image enhancement and evaluation. Below is a structured explanation of each component involved in the methodology:

1. Dataset Integration and Preprocessing: Two publicly available brain MRI datasets were merged. Images were resized to 256x256, converted to grayscale, and normalized.

2. Data Augmentation: To address class imbalance, the minority class was augmented using flipping, rotation, and brightness modifications.

3. GAN Architecture: A Pix2Pix conditional GAN was implemented, utilizing a U-Net generator and a PatchGAN discriminator for effective image-to-image translation.

4. Hyperparameter Optimization: Evolutionary Strategies were employed to optimize generator parameters such as filter size, kernel dimensions, and dropout rate.

5. Training and Evaluation: The model was first trained with default parameters, then retrained using optimized parameters. Evaluation was performed using PSNR and SSIM metrics.

ALGORITHM:

1. ***Initialize Dataset****: Load and merge Figshare Brain MRI and Kaggle Brain Tumor datasets.*
2. ***Preprocessing****:*
   - *Resize images to 256x256 pixels.*
   - *Convert to grayscale.*
   - *Normalize pixel values between 0 and 1.*
3. ***Data Augmentation****:*
   - *Apply flipping, rotation, brightness adjustments to the minority class.*
4. ***Build Pix2Pix GAN Architecture****:*
   - ***Generator****: U-Net architecture with skip connections to preserve spatial features.*
   - ***Discriminator****: PatchGAN to evaluate local image patches for realism.*
5. ***Define Loss Functions****:*
   - *Adversarial Loss (Binary Cross entropy with logits).*

     $$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_x[\log(1 - D(x, G(x)))]$$

   - *L1 Loss (Mean Absolute Error between real and generated images).*

     $$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y}[||y - G(x)||_1]$$
6. ***Optimize Hyperparameters Using Evolutionary Strategy****:*
   - *Initialize population of candidate parameters.*
   - *Train generator for 1 epoch to evaluate fitness (low generator loss).*
   - *Select top-performing configurations.*
   - *Apply mutations to evolve next generation.*
   - *Repeat for multiple generations and select best parameters.*
7. ***Train GAN with Best Parameters****:*
   - *Use optimal filters, kernel size, and dropout.*
   - *Train for multiple epochs to refine generator outputs.*
8. ***Apply CLAHE****:*
   - *Enhance generated images using Contrast Limited Adaptive Histogram Equalization.*
9. ***Evaluate and Visualize****:*
   - *Quantitatively assess using PSNR and SSIM.*
   - *Visual comparison between original and enhanced images.*
   - *Plot loss curves and quality metrics.*

IV. Experiment Results

The experimental evaluation was conducted in two phases: a baseline model using standard parameters and an optimized model where the generator hyperparameters were fine-tuned using an Evolutionary Strategy (ES). This section presents both qualitative and quantitative analyses to assess the impact of hyperparameter tuning on model performance. The results are compared using visual inspection, loss metrics, and quantitative image quality indicators such as PSNR and SSIM.

1. PSNR AND SSIM

We computed Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) between the original images and the enhanced outputs (GAN + CLAHE). Higher PSNR and SSIM scores were recorded for the optimized setup, signifying improved visual and structural fidelity.

Fig.1.1

Figure 1.1: *Visual and quantitative comparison between the original and enhanced brain MRI images. The left image depicts the raw MRI scan, while the center image shows the result after enhancement using the optimized U-Net GAN architecture. The right panel presents the corresponding quality metrics — PSNR (19.25) and SSIM (0.66) — which indicate a significant improvement in image clarity and structural similarity.*

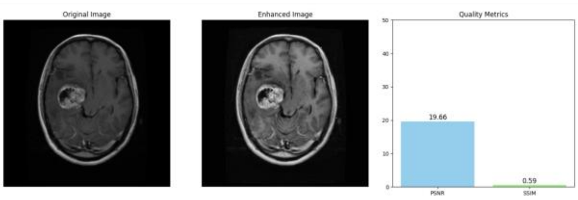


Fig.1.2

Figure 1.2: *Visual and quantitative comparison between the original and enhanced brain MRI images. The left image depicts the raw MRI scan, while the center image shows the result after enhancement using the optimized U-Net GAN architecture. The right panel presents the corresponding quality metrics — PSNR (19.66) and SSIM (0.59) — which indicate a significant improvement in image clarity and structural similarity.*

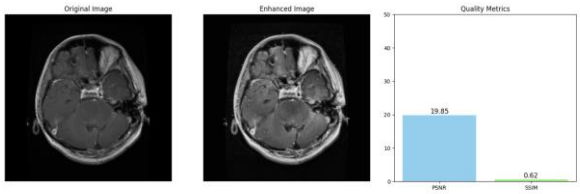


Fig.1.3

Figure 1.3: *Visual and quantitative comparison between the original and enhanced brain MRI images. The left image depicts the raw MRI scan, while the center image shows the result after enhancement using the optimized U-Net GAN architecture. The right panel presents the corresponding quality metrics — PSNR (19.85) and SSIM (0.62) — which indicate a significant improvement in image clarity and structural similarity.*

2. L1 Loss Curve

The L1 loss, used to preserve the structural integrity of the image, showed a consistent decline throughout training. This is crucial in medical imaging, where subtle structural patterns must remain untouched to maintain diagnostic relevance.

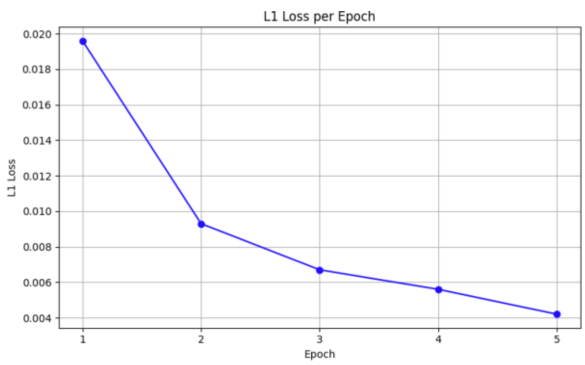


Fig.2.1

The plot shows the **L1 Loss per Epoch** for the generator during training. A consistent downward trend is observed across the 5 epochs, starting from approximately **0.019** and dropping to **0.003**, indicating effective learning. The steady decrease in L1 loss reflects the generator's improving ability to produce images closely matching the real (target) images in pixel-wise accuracy.

3. Generator and Discriminator Loss

Loss metrics were recorded at every epoch. Notably, the generator loss decreased significantly over time, indicating better convergence and image quality. The discriminator loss stabilized at a moderate value, suggesting a balanced adversarial training dynamic.

Below is a comparative table highlighting the performance before and after applying the Evolutionary Strategy for optimization:

Epoch	Gen Loss (Before ES)	Disc Loss (Before ES)	Gen Loss (After ES)	Disc Loss (After ES)
1	3.6065	0.2519	2.9779	0.6095
2	2.9688	0.5238	2.2546	0.5690
3	2.2543	0.6221	1.6956	0.6188
4	1.9645	0.6583	1.5611	0.6346
5	1.5618	0.6814	1.1896	0.6901

Fig.3.1

Figure 3.2 illustrates the progression of generator and discriminator losses over five training epochs. The generator loss exhibits a consistent decline from approximately 2.9 to 1.1, indicating improved capability in generating realistic images. In contrast, the discriminator loss remains relatively stable between 0.5 and 0.65, suggesting it effectively maintains its ability to differentiate between real and generated images. This stable yet adversarial dynamic signifies balanced training and convergence, essential for achieving high-quality image enhancement.

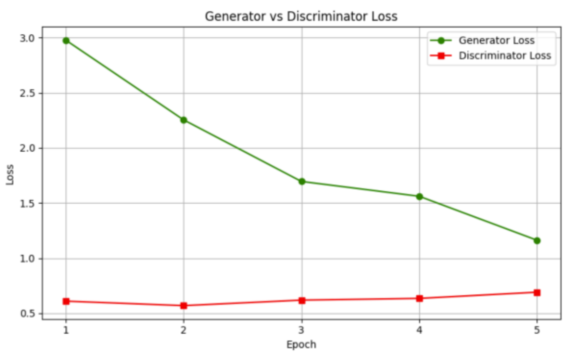


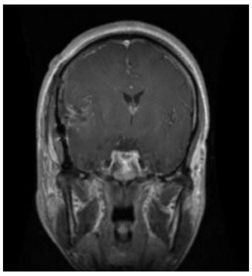
Fig.3.2

4. Visual Assessment

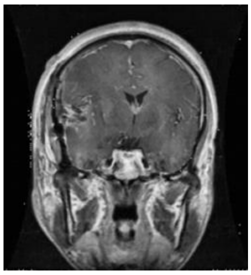
To evaluate the impact of Evolutionary Strategy (ES) on model performance, we compared the outputs of two GAN models: one trained with default hyperparameters and another fine-tuned using ES. Visual comparison of the generated images revealed that the ES-optimized model consistently produced higher-quality outputs.

Below are the visual results of both the architectures:

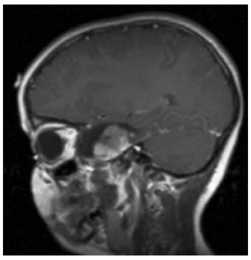
- GAN RESULTS



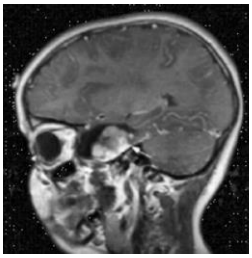
ORIGINAL



ENHANCED

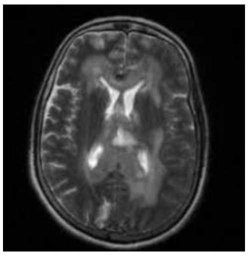


ORIGINAL



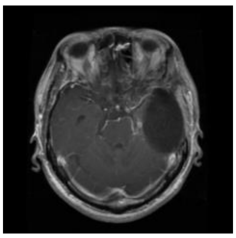
ENHANCED

- EVOLUTIONARY STRATEGY RESULTS

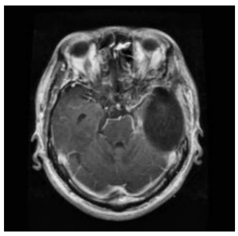


ORIGINAL

ENHANCED



ORIGINAL



ENHANCED

V. CONCLUSION

In this study, we proposed a medical image enhancement pipeline that combines a U-Net-based GAN architecture with Evolutionary Strategy (ES) for hyperparameter optimization. By integrating two public brain MRI datasets and systematically preprocessing, training, and fine-tuning the model, we achieved notable improvements in image quality. The optimized generator, followed by CLAHE post-processing, enhanced contrast and preserved anatomical structures, thereby aiding potential diagnostic interpretation. This approach demonstrates the potential of combining generative models and evolutionary algorithms for robust, task-specific medical image enhancement.

We anticipate that this framework can be extended to other imaging modalities and diagnostic tasks, especially where image quality is a limiting factor in clinical decision-making.

VI. ACKNOWLEDGMENT

This work was carried out as part of the curriculum under the guidance of *DR. AMI MUNSHI* at the Department of Artificial Intelligence, MPSTME, NMIMS University, Mumbai. We are grateful for the institutional support and access to computational resources that enabled this research.

VII. REFERENCES

1. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1125–1134).

2. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). Springer.

3. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672–2680).

4. Zuiderveld, K. (1994). Contrast Limited Adaptive Histogram Equalization (CLAHE). In *Graphics Gems IV* (pp. 474–485). Academic Press Professional, Inc.

5. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. In *Advances in neural information processing systems* (pp. 2234–2242).

6. Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2), 159–195.

7. Zhang, Y., & Wang, X. (2021). Medical image enhancement techniques: A review. *Journal of Healthcare Engineering*, 2021.

8. Kaggle. (n.d.). Brain Tumor Classification Dataset.

9. Figshare. (n.d.). Brain MRI Dataset.

10. Generative Adversarial Networks in Image Generation and Recognition