Elevating Diagnostic Accuracy: Advanced GAN-Enhanced High-Resolution Medical Imaging for Superior Disease Detection

Vathana D.[1], Babu S.[1,*]

[1]Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology Kattankulathur – 603203, Chengalpattu District, Tamil Nadu, India

Emails: vathanad@srmist.edu.in; babus@srmist.edu.in

Abstract

Advanced imaging in medical has become crucial in the early identify diseases because they reveal the important structural features of the human body. But it is almost impossible to get such high resolution images in real life situation due to the factors such as image capture and processing equipment, and environmental factors that affect the outcome of the image. This work proposes a sub-type of GAN that is used in enhancement of images particularly in medical fields. The generator of the Med-GAN extracts a high-resolution image from a low-resolution one with the help of novel features learned by the model. The approach of reconstructing high resolution from multiple parallel streams of lower resolution employs deconvolution algorithms with multiple scale fusions that produce better high resolution representations as compared to the technique of bilinear interpolation. The performances of the proposed Med-GAN are tested on two publicly available COVID-19 CT datasets and one private medical image dataset which shows that the proposed method outperforms the existing methods in performance comparisons. Consequently, for PSNR, the score improves from 24.103 dB corresponding to the Initial Approach of the “BRaTS (FLAIR)” dataset to 25.496 dB for the Proposed Method; whereas for SSIM the score increases from 0.782 to 0.812.se types of high-resolution images are usually impossible to get due to limits in imaging devices, environmental conditions, and human factors. This work proposes the Med-GAN: an Enhanced Super-Resolution Generative Adversarial Network tuned for medical image enhancement. The Med-GAN generator learns high-resolution representations from low-resolution images via advanced feature extraction methods. Deconvolution algorithms with multi-scale fusions recover better high-resolution representations from multiple parallel streams of lower resolutions in this approach compared to traditional bilinear interpolation methods. Evaluated on two publicly available COVID-19 CT datasets and one custom medical image dataset, the proposed Med-GAN significantly outperforms existing techniques in performance comparisons. In particular, PSNR rises from 24.103 dB for the "BRaTS (FLAIR)" dataset in the Initial Approach to 25.496 dB in the Proposed Method, while SSIM increases from 0.782 to 0.812. If that is the case then it could be said that the solution of the proposed Med-GAN is one of the most realistic means for improving the quality of medical images and therefore contributes to better diagnostics of diseases

Keywords: High-resolution medical imaging; Image augmentation; Enhanced Super-Resolution Generative Adversarial Networks; Med-GAN; Super-resolution

1. Overview

Typically, medical images are understood by professionals in the medical area, such as radiologists and physicians. However, the discrepancies among various specialists and the intricacies of medical imaging provide significant challenges for professionals involved in the precise and consistent detection of disorders. Computerized approaches facilitate the semi-automated or fully automatic completion of challenging image analysis tasks, allowing specialists to make diagnoses that are objective, efficient, and precise. Therefore, the exploration of deep learning techniques for the assessment of medical images has consistently been a fascinating field of study [1-3].

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

Deep Convolutional Neural Network (CNN) architectures have been extensively employed across various domains in the field of medical diagnostics. These structures are employed for picture segmentation, categorization, and authentication, and content-based image retrieval, assessment of illness severity, identification of cancers, lesions, and tissues [4]. Although CNNs have demonstrated superior performance compared to conventional machine learning techniques, there are still substantial challenges to address in order to enhance their ability to generalize and create robust models. In order to ensure robustness and great generalizability, it is imperative that the training data sets include a substantial number of Images that encompass all potential variations. Nevertheless, the quantity of medical images that are typically accessible is typically restricted. The issue of data scarcity can be caused by a variety of factors, including an insufficient number of patients for specific diseases, patients' reluctance to consent to the use of their images, insufficient availability of medical equipment, and the inability to obtain images that meet specific requirements. Acquiring images of patients from a diverse array of racial or geographical contexts may present challenges.

Labelling the acquired Images can be quite challenging, especially when compared to the straightforward process of labelling natural photographs. This remains a significant issue, even after a sufficient number of images have been collected. It requires specialized knowledge from medical professionals and is a task that requires a significant amount of time. In addition, there are legal and privacy concerns associated with labelling medical images. Addressing the issue of imbalanced data is crucial when developing deep network-based methods in order to avoid overfitting, skewed outcomes, and incorrect results. Image augmentation approaches tackle these challenges by using an equal number of images and automatically expanding training datasets. They require strong and advanced techniques to enhance their ability to learn and generalize, in order to create deep networks. However, the literature has utilized a range of different strategies to enhance the data with various types of images. The literature should clearly state whether the data augmentation strategy leads to more efficient outcomes for specific types of images. This ambiguity arises from the different types of illnesses, network topologies, and the amount of datasets used for training and testing. Recently, there have been significant advancements in super-resolution (SR) methods using CNNs, resulting in excellent performance [5-7]. The networks have evolved over time, starting with the original SRCNN and advancing to more sophisticated models such as Very Deep Super-Resolution, Deep Recursive Residual, Memory, and Residual Channel Attention Networks.

In addition, there are other effective techniques that involve connecting multiple feature extraction modules to create comprehensive networks. Some examples include the Residual Dense Network [8], Information Distillation Network [9], Multi-Scale Residual Network [10], and Super-Resolution Feedback Network [11]. Emphasizing the importance of the unique capabilities of each module. As mentioned in reference [12], the GAN model presents a fresh perspective on image generation and acts as a fundamental model for producing high-resolution images. The SRGAN was the first to incorporate the GAN framework into SR reconstruction, as demonstrated in [13]. This approach has effectively enhanced the visual quality of images and provided a more precise representation of high-frequency information. However, due to the straightforward architecture of the SRGAN generating network, the extracted features may not always be adequate, leading to a decrease in the quality of the reconstruction. Later, a number of innovative super-resolution (SR) techniques were introduced that utilized GAN models and deep convolutional networks to improve image quality in different ways [14-17]. Although GAN-based SR models can yield favourable outcomes, they still have a number of drawbacks: (i) the training process and super-resolution capability of the initial GAN architecture show a lot of volatility. (ii) The oversimplification of the generation network in the framework hampers its suitability for feature extraction in the super-resolution task. Consequently, the degree of reconstruction is greatly affected, leading to poor picture feature extraction. After careful evaluation of GANs and CNNs, we have developed a new architecture called Med-GAN. This method utilizes advanced techniques in representation learning to enhance the quality of medical images.

Here is the outline of the remaining sections: Section 2 discusses the work that is relevant to the topic. In Section 3, you will find a comprehensive overview of Med-GAN's data gathering, GAN architecture, and feature extraction processes. Section 4 provides an analysis of the findings, comparisons to current approaches, and assessments of the effects of aggregation. Section 5 provides an overview of the findings and offers suggestions for future research.

2. Related works

In the study conducted by [18], they were able to achieve improved tumour segmentation and classification accuracy by implementing additional noise and sharpening approaches. Another study focused on studying images that were enhanced using various methods such as blurring, noise addition, random cropping, translation, and rotation. By employing these techniques, the algorithm demonstrated a higher level of effectiveness in predicting ages, diagnosing schizophrenia, and classifying sexes [19]. Various studies have enhanced the accuracy of tumour segmentation by employing techniques such as elastic deformation, random scaling, and rotation [20]. Although

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

the literature primarily supports those procedures, synthetic images have also been used to achieve augmentation. Various methods have been utilized to generate synthetic images, including GAN and modified Mixup architectures, along with Mixup, which combines two randomly selected images with their respective labels. An example of a variant of Mixup, TensorMixup is a voxel-based method that combines two picture patches using a tensor. As demonstrated in [22], this approach has been utilized to enhance the accuracy of tumour segmentation.

A recent study utilized an augmentation strategy that relied on Generative Adversarial Networks (GANs) to perform cerebral vascular segmentation. Medical professionals rely on computerized technologies to aid in tasks such as segmenting lung tissue or nodules, detecting nodules, predicting malignancy levels, and diagnosing diseases like coronavirus (COVID-19). There are various techniques available to enhance CT lung imaging, such as rotating, flipping, and scaling the images. Image enhancement techniques were employed to enhance the automatic diagnosis of COVID-19 [24]. Utilizing additional enhancements alongside Gaussian noise, [25] improved the efficiency of COVID-19 diagnosis. In addition, several models of Generative Adversarial Networks (GAN) have been used to improve lung CT images by generating new synthetic images. When enhancing lung CT images with a Generative Adversarial Network (GAN), researchers usually concentrate on generating the particular area of the image that shows lung nodules. Given the small lung nodules, it would be more efficient to generate a scaled-down image or focus specifically on the target area instead of capturing a wide shot of the entire scene. 26 in Developed synthetic images depicting the region of lung nodules utilizing directed GAN architectures. Using nodule size information alongside guided GANs, researchers were able to generate images of lung nodules in different sizes.

Augmentation techniques are frequently utilized in mammography images to improve the automated identification, segmentation, and detection of abnormalities in the breast. Errors in detecting or identifying lesions can lead to unnecessary biopsies or incorrect diagnoses. Instead of applying augmentation techniques to the entire image, numerous studies have concentrated on specific positive and negative areas that have been extracted [28]. Positive patches are obtained by extracting regions marked by medical professionals, while negative patches are randomly selected from the same image. Prior to implementing an augmentation technique, it is highly advisable to refine and segment the data, ensuring that the segments are properly defined and concluded. This approach is the most effective strategy to guarantee outstanding performance in detecting and classifying the lesions. Various image refining techniques, including noise removal and resizing, were utilized. Later on, the implementation of segmentation and augmentation stages took place [29]. Prior to dividing the images into distinct sections and applying augmentation, any noise and artifacts were eliminated. Applying various techniques such as scaling, noise addition, replicating, shearing, and rotation is a common approach to enhance mammography images or specific regions within them.

Various augmentation techniques were employed to improve the automatic categorization of images into three categories: standard, malignant, and benign [29]. These techniques included mirroring, zooming, and resizing. In addition, these techniques have been seamlessly integrated with GANs to produce innovative augmented synthetic visuals. We have successfully integrated flipping with deep segmentation GAN to perform binary classifications of images, distinguishing between mass-laden and standard images. GAN-based augmentations have been extensively utilized. A prime illustration is the contextual GAN, which generates novel images by amalgamating injuries derived from the adjacent tissue. This technique to improve binary classification by effectively differentiating between normal and malignant cases. In a separate study, researchers successfully generated artificial lesion images that provided detailed information about the lesion's border and texture. This was achieved by combining contextual and deep convolutional GAN methods. Deep learning algorithms currently utilize these properties for a range of applications, such as detecting glaucoma, classifying diabetic retinopathy (DR), and assessing the severity of DR. Eye fundus images are often enhanced using various augmentation techniques such as movement, shearing, flipping, and translation. These techniques are employed to enhance the performance of deep learning algorithms. In a recent study by [32], a comprehensive method was employed to diagnose glaucoma. This method involved rotation, shearing, zooming, flipping, and shifting.

[33] A technique of enhancement was employed, involving the application of blurring and shifting. This technique was utilized prior to conducting binary classification (healthy vs. non-healthy) and multi-class classification (5 classes) of images. The goal was to differentiate between different stages of DR. Another study [34] utilized a combination of movement, translation, mirroring, and zooming techniques to automatically detect malaria retinopathy. In addition, pixel-intensity data has been utilized in specific augmentation methods due to the coloured nature of fundus images. A technique known as channel-wise random gamma correction has been utilized to improve the segmentation of retinal arteries [35]. The authors have successfully implemented a channel-wise vascular enlargement through the use of morphological modifications. In a separate study, researchers [36] employed various picture augmentation techniques such as zooming and Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance the robustness of DR identification. In addition, the utilization of GAN models

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

has greatly aided in the generation of synthetic images, thereby improving augmentations. Researchers have used conditional GAN [37] and deep convolutional GAN [38] augmentations to improve the performance of algorithms for automatically categorizing diabetic retinopathy (DR) based on its severity level. In certain cases, new images have been generated and improved with retinal characteristics, including vascular and lesion features. For instance, researchers in [39] study (2020) utilized heuristic image augmentation to create structures similar to NeoVessel (NV). This method was used to improve the detection of proliferative diabetic retinopathy (DR), a stage of DR where new blood vessels form. Considering the desired appearance and positioning of the NVs, different images can be created for this improvement by utilizing various NVs such as trees, wheels, and brooms.

3. Equipment and Techniques

A. Information collection

Brain magnetic resonance imaging (MRI) was included in the BRaTS Multimodal Brain Tumor Images Segmentation Challenge (40), which was only recently finished. FLAIR, T2, T1-CE, and T1-weighted are the four MR imaging models that are included in the data for each individual patient. Numerous modal images each give a unique set of information on tumours. This item involves the utilization of FLAIR photographs. The BRaTS collection contains a total of 285 photographs, 210 of which are of HGG cases and 75 of which are of LGG cases. There are photographs of the lungs that may be found in the LIDC-IDRI database [41–42] as well as in the LUNA16 challenge (17). A total of 1018 CT scans from 1010 different people are included in the database. These scans were analysed and interpreted by a group of four radiologists comprising an expert panel. Radiologists have categorised nodules into three distinct groups: "non-nodule ≥ 3 millimetres," "nodule < 3 millimetres," and "nodule ≥ 3 millimetres." Because of the utility of the images, those with a slice thickness greater than three millimetres were removed [44-45].

Other images that were eliminated included those that had slices that were either missing or had an inconsistent spacing between them. There were a total of 1018 cases, and 888 Images were removed from them. Mammography images of the breast are available through the INbreast database [46]. Breast masses are depicted in only 107 of the 410 Images, which results in 115 mass lesions. This is due to the fact that photographs may contain many lesions. There are 63 more malignant masses than there are benign ones, which total 52. According to [47], the Messidor database has 1200 eye fundus images that have been categorized into three danger levels: 0 risk, 1 risk, and 2 risk. These risk levels are determined by the risk of diabetic macular oedema. There are 974, 75, and 151 images in each of the categories.

B. Methodology

The objective of this research is to generate a high-resolution medical image by utilizing the bicubic operation on a low-resolution image, based on a high-resolution reference image. Consider two images, one with a low resolution (X) and the other with high quality (Y). The mapping function Gθ(•) represents a direct connection between the starting point and the endpoint. The solutions to the aforementioned problems have the potential to produce X and Y.

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} L(G_{\theta}(X_i), Y_i), \quad (1)$$

The variable θ is employed in this particular context to denote the set of network parameters that require optimization. The function L(.) serves the purpose of functioning as a loss in order to minimize the discrepancy between X and Y. The quantity of drill samples is represented by the variable N.

An adversarial network, more specifically a general adversarial network (as shown in Figure 1[25]), is trained to engage in a minimax game. The objective of the generator network is to maximize a function, whereas the discriminator network aims to minimize it, in comparison. The competitive nature of this dynamic is the reason behind the term "adversarial." Generative Adversarial Networks (GANs) are trained by optimizing a loss function during their training process.

$$\begin{aligned} \min_{G} \max_{D} V(D, G) = {} & \mathbb{E}_{x \sim p\ \text{data}\ (x)}[\log D(x)] \\ & + \mathbb{E}_{z \sim pz(z)}[\log(1 - D(G(z)))]. \end{aligned} \quad (2)$$

For the purpose of addressing the generator network, which is characterized by the parameters θG, the term G(z) is utilized in this context. A random variable z is provided to the system, and it is taken from a prior distribution using the pz distribution. In order to accomplish its goal, G must first convert the variable z into a variable x that is derived from the distribution set of data. Using parameters θD, a distinct network known as D, which is referred to as the discriminator, is trained in order to achieve this goal. The primary objective of this analysis is to

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

differentiate between the genuine samples x ∼ data derived from a certain dataset and the fabricated samples ˆx ∼ pθG (x|z) that are generated by the generator. By employing this tactic, the generator is compelled to produce samples that are more compelling in order to deceive the discriminator into believing that they are genuine.

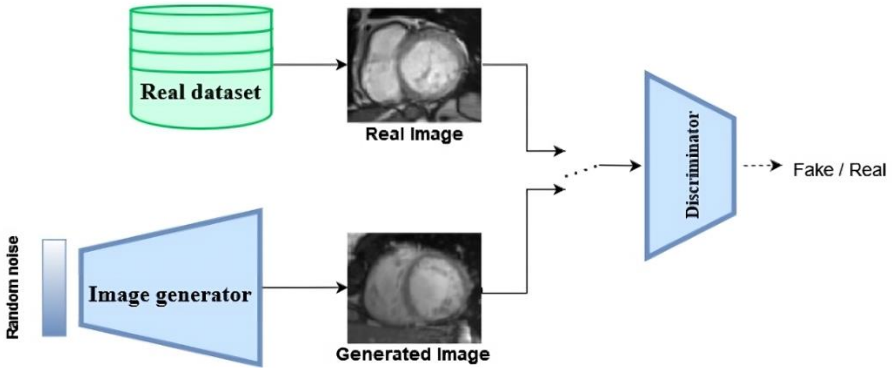

Figure 1. Architecture of conventional GAN [25]

Increasing the quality of the recovered image in GAN is primarily accomplished through the utilization of two modifications that are directed toward the structure of generator G: First, remove all of the BN levels from the structure. The second step is to replace the previous basic block with the RRDB, which is a combination of a multi-level residual network and dense connections (for an illustration, see Figure 3).

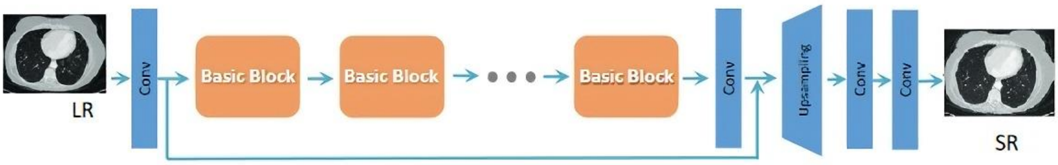

Figure 2. Architecture of SRGAN [48]

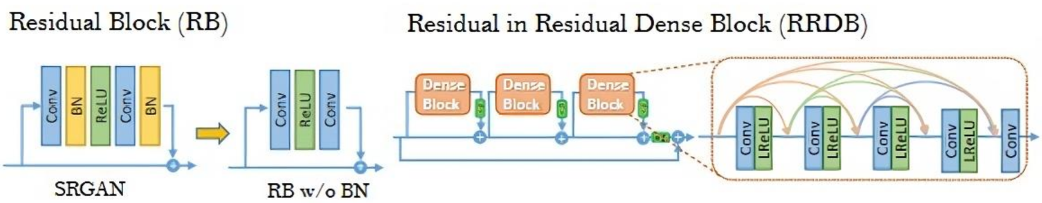

Figure 3. RRDB Structure [49]

In a number of applications that are centered on PSNR (Peak Signal-to-Noise Ratio), such as SR (Super-Resolution) and deburring, it has been demonstrated that excluding BN (Batch Normalization) layers can improve performance and reduce the amount of computing complexity involved. Batch normalization (BN) layers make advantage of the batch's mean and variance in order to standardize the features prior to training. Nevertheless, for testing purposes, we make use of the mean and variance that were calculated for the entire training dataset. It is possible that the employment of BN (Batch Normalization) layers could result in the appearance of undesired artifacts and limit the ability to generalize when there is a large variation in the statistics of the training dataset and the testing dataset. It has been demonstrated through empirical data that certain BN layers are more likely to introduce artifacts when a deep neural network is trained using a GAN architecture. The criterion for consistent performance throughout the training process is violated by the fact that these artefacts reveal themselves in a sporadic manner between iterations and in a variety of circumstances.

In order to ensure reliable training and consistent performance, the BN layers are removed. Furthermore, the elimination of BN (Batch Normalization) layers enhances the ability to generalize while also decreasing the computational complexity and memory usage. The newly introduced basic block in SRGAN is referred to as RRDB. Figure 4 provides a visual representation of this block. Figure 2 illustrates the high-level architecture

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

design along with its various components. Based on the available data, it has been observed that increasing the number of layers and connections consistently enhances performance. The recommended RRDB (Residual in Residual Dense Block) utilizes an architecture that is characterized by increased complexity and depth compared to the original residual block employed in SRGAN (Super-Resolution Generative Adversarial Network). The presented RRDB includes a structured implementation of residual learning on multiple layers. This design incorporates residuals within residuals, as depicted in Figure 3. The utilization of a multi-level residual network topology has been proposed. However, our RRDB stands out from other systems by incorporating a thick block into the primary channel, similar to the way it is done in. The utilization of dense connections is a design option that maximizes the capacity of the network.

The structure of both the discriminator and the generator has been enhanced by implementing the Relativistic GAN method. A relativistic discriminator is a tool used to determine the likelihood that an input image x is genuine and authentic, as opposed to a manufactured one xf. In contrast to the typical discriminator D in SRGAN, the relativistic discriminator focuses on assessing the possibility of an actual picture being more realistic than a manufactured one xr.

C. Loss functions

The term "measure of overall loss" refers to a quantitative assessment of the total amount of loss experienced in a given situation or context. The proposed model's LSR, following the SRGAN approach, can be described as a biased sum of individual loss functions.

$$L_{\text{Total}}^{\text{SR}} = a_1 L_{\text{MSE}}^{\text{SR}} + a_2 L_{\text{Gen}}^{\text{SR}} \tag{3}$$

$$L_{\text{MSE}}^{\text{SR}} = \frac{1}{r^2WH} \sum_{i=1}^{rW} \sum_{j=1}^{rH} \left(Y_{i,j} - G_{\theta_G}(X)_{i,j}\right)^2 \tag{4}$$

$$L_{\text{Gen}}^{\text{SR}} = \sum_{n=1}^{N} - \log D_{\theta_p}\left(G_{\theta_G}(X)\right) \tag{5}$$

The weighting parameters a1 and a2 are used to represent the importance of different factors in the optimization process for image super-resolution. The primary optimization objective, known as LSR MSE, is a measure of content loss. This objective is widely used and relied upon by many state-of-the-art methods in the field. The term "Localized Saliency Regularization" is an acronym that represents the concept it describes. The term "Gen" in the context of generative network antagonistic loss refers to the generator network's objective of deceiving the discriminator network. The variable "r" is used to represent the down-sampling factor in the down-sampling process. The width of the image is denoted by the variable "W", while its height is represented by the variable "H".

4. Results and Discussion

A. Execution Details

The medical Image database was populated with a variety of datasets encompassing different medical imaging modalities. The FLAIR images used in this study were obtained from the BRaTS database, which consists of a total of 285 images. A total of 210 instances are categorized as high-grade gliomas (HGG), while 75 cases are categorized as low-grade gliomas (LGG). The study utilized lung CT images obtained from the LIDC-IDRI and LUNA16 databases. The dataset included a total of 888 images collected from 1010 individuals. In addition, breast mammography images from the INbreast database were utilized. This database consists of 107 images that depict breast masses. The collection of images contains a total of 115 mass lesions, consisting of 52 benign and 63 malignant cases. Finally, eye fundus images from the Messidor database were utilized. The database consists of 1200 images that have been categorized into three risk levels for diabetic macular oedema. The augmentation technique was also implemented on the CT_349 and CT_19 datasets. A training dataset consisting of 700 images was generated to augment the data. This was achieved by rotating the original images by 90°, 180°, and 270°, and subsequently horizontally flipping them. The training process of the model was initiated using the Adam optimizer. A uniform learning rate of 0.0001 was applied to all layers. After every 50 epochs, the learning rate was reduced by half. The model achieved convergence after 200 iterations. The training process, which utilized a single Tesla P40 GPU, took approximately 9 hours.

B. Med-GAN versus State-of-the-Art: A Comparative Study

With the help of a medical imaging database that contains datasets from the BRaTS (FLAIR), LIDC-IDRI & LUNA16, INbreast, Messidor, and COVID-19 databases, the performance of the technique that has been suggested is tested. In order to guarantee a direct comparison, the approaches that were utilized in this study made use of the same training set as well as public model specifications.

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

The table 1 presents a comparison of the Peak Signal-to-Noise Ratio (PSNR) values obtained from different feature aggregation techniques applied to four medical imaging datasets: BRaTS (FLAIR), LIDC-IDRI & LUNA16, INbreast, and Messidor. The Peak Signal-to-Noise Ratio (PSNR) values serve as an indicator of the quality of the reconstructed images. Higher PSNR values correspond to better image quality. The initial approach produces the lowest Peak Signal-to-Noise Ratio (PSNR) values for all datasets: 24.103 for BRaTS, 26.127 for LIDC-IDRI & LUNA16, 27.005 for INbreast, and 26.732 for Messidor. This indicates that this method is less successful in preserving image quality. The standard method demonstrates an improvement over the initial approach by achieving higher Peak Signal-to-Noise Ratio (PSNR) values. Specifically, the PSNR values for BRaTS, LIDC-IDRI & LUNA16, INbreast, and Messidor are 25.042, 26.879, 27.874, and 27.223, respectively. These results indicate that the standard method enhances image quality by effectively aggregating features.

Enhanced performance is observed when additional techniques are integrated. The combination of the standard method and BI MR Fusion results in slightly improved Peak Signal-to-Noise Ratio (PSNR) values. Specifically, the PSNR values are 25.084 for BRaTS, 26.950 for LIDC-IDRI & LUNA16, 27.998 for INbreast, and 27.287 for Messidor. These values suggest a minor enhancement in image quality. The inclusion of deconvolutions in the standard method leads to a further improvement in PSNR values. Specifically, for BRaTS, the PSNR value increases to 25.112, for LIDC-IDRI & LUNA16 it increases to 27.082, for INbreast it increases to 28.094, and for Messidor it increases to 27.331. These results indicate that the feature aggregation process has been enhanced. The proposed method is a combination of elements from the initial approach, the standard method, and deconvolutions. It has been found to achieve the highest Peak Signal-to-Noise Ratio (PSNR) values across all datasets. Specifically, the PSNR values are 25.496 for BRaTS, 27.465 for LIDC-IDRI & LUNA16, 28.416 for INbreast, and 27.634 for Messidor. The aforementioned statement suggests that the proposed technique is highly efficient in maintaining the quality of images, thereby showcasing its superior ability to aggregate features across various medical imaging datasets.

Table 1: Comparison PSNR for Different Feature Aggregation Techniques across Medical Imaging Datasets

Techniques	BRaTS (FLAIR)	LIDC-IDRI & LUNA16	INbreast	Messidor
Initial Approach [13]	24.103	26.127	27.005	26.732
Standard Method [50]	25.042	26.879	27.874	27.223
Standard Method + BI MR Fusion [51]	25.084	26.950	27.998	27.287
Standard Method + Deconvolutions	25.112	27.082	28.094	27.331
Initial + Standard + Deconvolutions (Proposed)	25.496	27.465	28.416	27.634

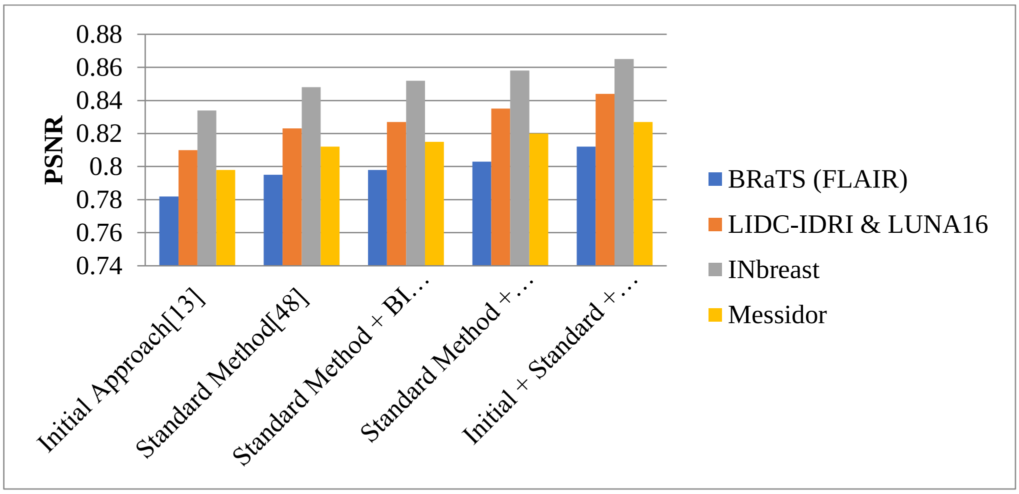

Figure 4. PSNR comparison

Four separate medical imaging datasets—BRaTS (FLAIR), LIDC-IDRI & LUNA16, INbreast, and Messidor—had their Structural Similarity Index (SSIM) values displayed in the following table. The feature aggregation procedures used in these datasets were varied. One quantitative measure for gauging how similar two images are

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

structurally is the Structural Similarity Index (SSIM). Higher values indicate better quality and more resemblance, and it gives a numerical value between 0 and 1. Elevated SSIM values signify the maintenance of picture quality and structural integrity. Among all datasets, the Initial Approach [13] shows the lowest SSIM values: 0.782 for BRaTS, 0.810 for LIDC-IDRI & LUNA16, 0.834 for INbreast, and 0.798 for Messidor. From what we can see, this method yields the worst outcomes in terms of both picture reconstruction accuracy and quality preservation. By attaining somewhat higher SSIM values (0.795, 0.823, 0.848, and 0.812, respectively), the Standard Method [50] shows improvements, signifying better picture quality and structural similarity. By combining the Standard Method with BI MR Fusion, we find even more improvement, leading to marginally greater SSIM values [49]. The SSIM values for BRaTS, LIDC-IDRI & LUNA16, INbreast, and Messidor are 0.798, 0.827, 0.852, and 0.815, respectively. The feature aggregation process is improved with the integration of BI MR Fusion, leading to better image quality. By utilizing a more sophisticated approach that improves the preservation of picture details and structural accuracy, the **Standard Method + Deconvolutions** shows an improvement in SSIM values (0.803, 0.835, 0.858, and 0.820).

The highest SSIM values across diverse datasets are demonstrated by the Proposed Method, which combines components from the first approach, standard method, and deconvolutions. Regarding BRaTS, LIDC-IDRI & LUNA16, INbreast, and Messidor, the SSIM values that the Proposed Method attains are 0.812, 0.844, 0.865, and 0.827, respectively. Out of all the approaches that were tested, this study shows which one is the best at aggregating features. Using this method consistently produces the most structurally similar and high-quality images across all datasets.

Table 2: Comparison of SSIM for Different Feature Aggregation Techniques across Medical Imaging Datasets

Techniques	BRaTS (FLAIR)	LIDC-IDRI & LUNA16	INbreast	Messidor
Initial Approach[13]	0.782	0.810	0.834	0.798
Standard Method[50]	0.795	0.823	0.848	0.812
Standard Method + BI MR Fusion[51]	0.798	0.827	0.852	0.815
Standard Method + Deconvolutions	0.803	0.835	0.858	0.820
Initial + Standard + Deconvolutions (Proposed)	0.812	0.844	0.865	0.827

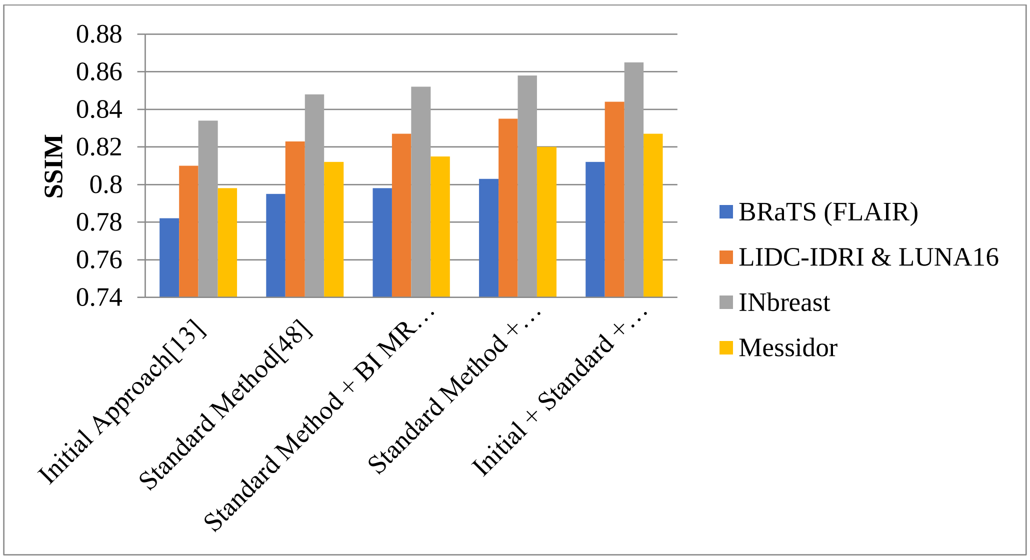

Figure 5. SSIM Comparison

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

C. Aggregation Impact Evaluation

In this section, we will evaluate the performance of the attribute aggregation component by subjecting the newly constructed medical picture database to various tests. The performance of our proposed feature aggregation component is now superior due to the incorporation of a deconvolution operation, when compared to SRGAN [13] and HRNet [50]. The table below displays the contrast (scale: 8×) of SSIM for various feature aggregation components. Our proposed technique demonstrates superior performance in terms of SSIM across multiple datasets including BRaTS (FLAIR), LIDC-IDRI & LUNA16, INbreast, and Messidor. The data can be visualized in Tables 1 and 2, as well as Figures 4 and 5. In our study, the term "BI" is employed to denote the up sampled bilinear interpolation process, while the term "MR" is utilized to represent multi-resolution.

5. Conclusion and Future Scope

We introduce a medical Image super-resolution network that utilizes a Generative Adversarial Network (GAN) and focuses on high-resolution representation learning. It optimally uses the characteristics of medical images to enhance the performance of SR, taking into account the benefits of Med-GAN, which can significantly improve the visual quality of the Images. To maintain HR representations and conduct repeated multi-scale fusions, it is essential to use HRNet as the SR foundation, consequently improving HR representations to enable SR. Deconvolution processes are also taken advantage of to bring back high-resolution (HR) reconstructions from low-resolution (LR) images. The objective is to generate more comprehensive aggregated features, surpassing the existing highest development methods in restoration with LR images. In the future, our focus will be on studying advanced multi-scale transform methods that incorporate the SR job to utilize characteristics extracted from medical Images used in HRNetV2 effectively. Quantitative and qualitative proof proves the strategy is effective in the experimental outcomes.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] Tsuneki M (2022) Deep learning models in medical image analysis. J Oral Biosci. 2022 Sep;64(3):312-320.

[2] Van der Velden BH, Kuijf HJ, Gilhuijs KG, Viergever MA (2022) Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal.

[3] Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, Mannel RS, Liu H, Zheng B, Qiu Y (2022a) Recent advances and clinical applications of deep learning in medical image analysis. Med Image Anal.

[4] Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ (2021) Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. Neurocomputing 444:92–110.

[5] 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), pp. 184–199, Springer, Zurich, Switzerland, September 2014.

[6] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using intense convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1646–1654, IEEE, San Juan, PR, USA, June 2016.

[7] Y. Tai, J. Yang, and X. M. Liu, "Image super-resolution via a deep recursive residual network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3147–3155, IEEE, Honolulu, HI, USA, July 2017.

[8] Y. L. Zhang, Y. P. Tian, Y. Kong, and B. N. Zhong, “Residual dense network for image super-resolution,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

[9] Z. Hui, X. M. Wang, and X. B. Gao, “Fast and accurate single image super-resolution via information distillation network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 723–731, IEEE, Salt Lake City, UT, USA, June 2018.

[10] J. Li, F. Fang, K. Mei, and G. Zhang, “Multi-scale residual network for image super-resolution,” in Proceedings of the European Conference on Computer Vision (ECCV), pp. 527–542, Springer, Munich, Germany, September 2018.

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

[11] Z. Li, J. L. Yang, Z. Liu, X. M. Yang, G. Jeon, and W. Wu, "Feedback network for image super-resolution," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3867–3876, IEEE, Long Beach, CA,USA, June 2019.

[12] I. J. Goodfellow, J. P. Abadie, and M. Mirza, "Generative adversarial nets," in Proceedings of the Annual Conference on Neural Information Processing Systems (ICONIP), pp. 2672–2680, Curran Associates, Montreal, Canada, December 2014.

[13] C. Ledig, L. Theis, F. Husz'ar, A. Cunningham, and A. Acosta, "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 105–114, IEEE, Honolulu, HI, USA, July 2017

[14] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, https://arxiv.org/abs/1701.07875.

[15] P. Isola, J. Y. Zhu, and T. Zhou, "Image-to-image translation with conditional adversarial networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134, IEEE, Honolulu, HI, USA, July 2017.

[16] H. Gao, Z. Chen, B. Huang, J. Chen, and Z. Li, "Image super-resolution based on conditional generative adversarial network," IET Image Processing, vol. 14, no. 13, pp. 3006–3013,2020

[17] X. L. Zun, H. J. Zhong, and L. R. Xing, "Multi-scale generative adversarial networks for image super-resolution algorithms," Science Technology and Engineering, vol. 20, no. 13, pp. 5217–5223, 2020.

[18] Khan AR, Khan S, Harouni M, Abbasi R, Iqbal S, Mehmood Z (2021) Brain tumour segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. Microsc Res Tech 84:1389–1399

[19] Dufumier B, Gori P, Battaglia I, Victor J, Grigis A, Duchesnay E (2021) Benchmarking CNN on 3d anatomical brain MRI: architectures, data augmentation and deep ensemble learning. arXiv preprint, pp 1–25.

[20] Isensee F, J¨ager PF, Full PM et al (2020) nnu-Net for brain tumor segmentation. In: International MICCAI brainlesion workshop. Springer, Cham, pp 118–132

[21] Fidon L, Ourselin S, Vercauteren T (2020) Generalized wasserstein dice score, distributionally robust deep learning, and ranger for brain tumor segmentation: brats 2020 challenge. in: International MICCAI brain lesion workshop, Lima, Peru, pp 200–214

[22] Wang Y, Ji Y, Xiao H (2022) A Data Augmentation Method for Fully Automatic Brain Tumor Segmentation. arXiv preprint, pp 1–15

[23] Kossen T, Subramaniam P, Madai VI, Hennemuth A, Hildebrand K, Hilbert A, Sobesky J, Livne M, Galinovic I, Khalil AA, Fiebach JB (2021) Synthesizing anonymized and labelled TOF-MRA patches for brain vessel segmentation using generative adversarial networks. Comput Biol Med 131:1–9

[24] Hu R, Ruan G, Xiang S, Huang M, Liang Q, Li J (2020) Automated diagnosis of COVID-19 using deep learning and data augmentation on chest CT. medRxiv, pp 1–11

[25] Alshazly H, Linse C, Barth E et al (2021) Explainable covid-19 detection using chest ct scans and deep learning. Sensors 21:1–22

[26] Wang Q, Zhang X, Zhang W, Gao M, Huang S, Wang J, Zhang J, Yang D, Liu C (2021) Realistic lung nodule synthesis with multi-target co-guided adversarial mechanism. IEEE Trans Med Imaging 40:2343–2353

[27] Nishio M, Muramatsu C, Noguchi S, Nakai H, Fujimoto K, Sakamoto R, Fujita H (2020) Attribute-guided image generation of three-dimensional computed tomography images of lung nodules using a generative adversarial network. Comput Biol Med

[28] Karthiga R, Narasimhan K, Amirtharajan R (2022) Diagnosis of breast cancer for modern mammography using artificial intelligence. Math Comput Simul 202:316–330

[29] Zeiser FA, da Costa CA, Zonta T et al. (2020) Segmentation of masses on mammograms using data augmentation and deep learning. J Digit Imaging 33:858–868

[30] Alyafi B, Diaz O, Marti R (2020) DCGANs for realistic breast mass augmentation in X-ray mammography. IN: Medical imaging 2020: computer-aided diagnosis, International Society for Optics and Photonics, pp 1–4

[31] Shen T, Hao K, Gou C, Wang FY (2021) Mass image synthesis in mammogram with contextual information based on GANS. Comput Methods Programs Biomed

[32] Shyamalee T, Meedeniya (2022) D CNN based fundus images classification for glaucoma identification. In: 2nd International conference on advanced research in computing (ICARC), Belihuloya, Sri Lanka, pp 200–205

[33] Tufail AB, Ullah I, Khan WU, Asif M, Ahmad I, Ma YK, Khan R, Ali M (2021) Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with a limited number of samples. Wirel Commun Mob Comput 2021:1–15

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024

[34] Kurup A, Soliz P, Nemeth S, Joshi V (2020) Automated detection of malarial retinopathy using transfer learning. In: IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Albuquerque, USA, pp 18–21

[35] Sun X, Fang H, Yang Y et al. (2021) Robust retinal vessel segmentation from a data augmentation perspective. In: International workshop on ophthalmic medical image analysis, pp 189–198

[36] Agustin T, Utami E, Al Fatta H (2020) Implementation of data augmentation to improve performance CNN method for detecting diabetic retinopathy. In: 3rd International Conference on information and Communications Technology (ICOIACT), Indonesia, Yogyakarta, pp 83–88

[37] Zhou Y, Wang B, He X, Cui S, Shao L (2020) DR-GAN: conditional generative adversarial network for fine-grained lesion synthesis on diabetic retinopathy images. IEEE J Biomedical Health Inf 26:56–66

[38] Balasubramanian R, Sowmya V, Gopalakrishnan EA, Menon VK, Variyar VS, Soman KP (2020) Analysis of adversarial based augmentation for diabetic retinopathy disease grading. In: 11th International Conference on Computing, Communication and Networking Technologies (ICT), India, Kharagpur, pp 1–5

[39] Araújo T, Aresta G, Mendonça L et al. (2020) Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images. IEEE Access 8:462–474

[40] Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara RT, Berger C, Ha SM, Rozycki M, Prastawa M (2018) Identifying the best machine learning algorithms for brain tumour segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint, pp 1–49

[41] Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, Zhao B, Aberle DR, Henschke CI, Hoffman EA, Kazerooni EA (2011) The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38:915–931

[42] Wang W, Luo J, Yang X, Lin H (2015) Data analysis of the lung imaging database consortium and image database resource initiative. Acad Radiol 22:488–495

[43] Setio AA, Traverso A, De Bel T, Berens MS, Van Den Bogaard C, Cerello P, Chen H, Dou Q, Fantacci ME, Geurts B, van der Gugten R (2017) Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Med Image Anal 42:1–3

[44] Manos D, Seely JM, Taylor J, Borgaonkar J, Roberts HC, Mayo JR (2014) The lung reporting and data system (LU-RADS): a proposal for computed tomography screening. Can Assoc Radiol J 65:121–134

[45] Naidich DP, Bankier AA, MacMahon H, Schaefer-Prokop CM, Pistolesi M, Goo JM, Macchiarini P, Crapo JD, Herold CJ, Austin JH, Travis WD (2013) Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. Radiology 266:304–317

[46] Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS (2012) Inbreast: toward a full-field digital mammographic database. Acad Radiol 19:236–248

[47] Decencière E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, Gain P, Ordonez R, Massin P, Erginay A et al (2014) Feedback on a publicly distributed image database: the messidor database. Image Anal Stereol 33:231–234.

[48] Ledig, C., Theis, L., Husz´ar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image superresolution using a generative adversarial network. In: CVPR (2017)

[49] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: CVPR (2018)

[50] S. H. Gao, M. M. Cheng, K. Zhao, X. Y. Zhang, M. H. Yang, and P. Torr, “Res2Net: a new multi-scale backbone architecture," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), IEEE, vol. 43, 2020

[51] K.Sun, B. Xiao, D. Liu, and J. D. Wang, “Deep high-resolution representation learning for human pose estimation,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2019

DOI: https://doi.org/10.54216/FPA.170214
Received: January 29, 2024 Revised: April 28, 2024 Accepted: September 27, 2024