

3D Magnetic Resonance Image Denoising using Wasserstein Generative Adversarial Network with Residual Encoder-Decoders and Variant Loss Functions

Hanaa A. Sayed¹, Anoud A. Mahmoud^{2*}, Sara S. Mohamed³

Dept. of Computer Science-College of Computer Science and Engineering, Taibah University, KSA-Dept. of Mathematics, Faculty of Computers and Information, Assiut University, 71516 Assiut, Egypt¹
Dept. of Mathematics and Computer Science-Faculty of Science, New Valley University, 72511 New Valley, Egypt^{2,3}

Abstract—Magnetic resonance imaging (MRI) is frequently contaminated by noise during scanning and transmission of images, this deteriorates the accuracy of quantitative measures from the data and limits disease diagnosis by doctors or a computerized system. It is common for MRI to suffer from noise commonly referred to as Rician noise because the uncorrelated Gaussian noise is present in both the real and imaginary parts of a complex K-space image with zero mean and equal standard deviation, the distribution of noise in magnitude MR images typically tends to be related to Rician distributions. To remove the Rician noise from an MRI scan, deep learning has been used in the MRI denoising method to achieve improved performance. The proposed models were inspired by the Residual Encoder-Decoder Wasserstein Generative Adversarial Network (RED-WGAN). Specifically, the generator network is residual autoencoders combined with the convolution and deconvolution operations, and the discriminator network is convolutional layers. As a result of replacing Mean Square Error (MSE) in RED-WGAN with Structurally Sensitive Loss (SSL), RED-WGAN-SSL has been proposed to overcome the loss of important structural details that occurs because of over-smoothing the edges. The RED-WGAN-SSIM model has also been developed using Structural Similarity Loss SSIM. The proposed RED-WGAN-SSL and RED-WGAN-SSIM models are formed by using the SSL, SSIM, Visual Geometry Group (VGG), and adversarial loss that are incorporated to form the new loss function. They preserved the informative details and fine image better than RED-WGAN, so our models could effectively reduce noise and suppress artifacts.

Keywords—Deep learning; image denoising; MRI; Wasserstein GAN; loss function

I. INTRODUCTION

MRI is a medical imaging process that produces multidimensional images of the inside of the body; it uses powerful magnets and radio waves generated by computers rather than injecting contrast agents. It is considered one of the most attractive modalities that have been used in the diagnosis and treatment of several neurological diseases because it can show 3D details of internal living tissues and the human body organs. MRI plays an increasingly important role in pathological and physiological diagnostics and scientific

research. Physiological noise impedes the acquisition of signals and contaminates raw data sets by artificial outliers. As a result of this practical issue, more advanced technologies have difficulty being applied in clinical research. Increasing noise levels may have a bad effect not only on the accuracy of computed diagnostic systems, but also on manual disease inspection and the reliability of quantitative image processing including segmentation, registration, visualization, super-resolution, and classification [1]. Raw data is usually polluted by White Additive Gaussian Noise (WAGN) in the real and imaginary parts. This noise is assumed to have equal variance and zero mean in the entire K space of the data, meaning that it affects both the real and imaginary parts of the data equally. Rician noise, on the other hand, is signal-dependent, which makes it harder to separate from the signal and can result in biased estimates of image intensity. Additionally, in high SNR regions, the Rician noise is close to the Gaussian distribution. To achieve reliable analysis results, it is necessary to remove noise before performing further image processing.

In MRI denoising, the goal is to effectively restore a clean image from a contaminated MR image and preserve valuable information [2]. In the past, many research attempts for MRI denoising were made to remove additive noise, most of which used the Rician noise model. In general, these methods can be categorized into three types: spatial filtering, transform domain filtering, and statistical methods [3]. The spatial domain techniques are directly applied to image pixels [4]. There are several traditional spatial image filters, including median [5], Gaussian [6], Wiener [7], diffusion [8], and bilateral filters. Anisotropic diffusion filter [9] significantly retained informative details of edges and reduced the noise from the images by smoothing local regions in the image, but the image was still blurry. This filter tried to avoid blurring of the edges by utilizing the edge-stopping function. A transform domain image filter is different from a spatial domain image filter. In that transform domain filtering methods first transform the space domain into another domain, and then they process the transformed image in the new domain based on the different characteristics of the image and its noise such as the frequency and wavelet domains [10]. Rician noise in MRI data has been successfully denoised by well-known block matching 3D

(BM3D) [11]. Higher-order singular value decomposition (HOSVD) [12] was developed to denoise MR volume data, and its performance was improved compared to BM3D.

Deep Learning (DL) made impressive progress in image processing and computer vision fields by introducing new effective methodologies. It has been used on low-level tasks to denoise [13,14], deblur [15], and restore super-resolution images [16]. CNNs and autoencoders achieved competitive performance with state-of-the-art methods, such as BM3D and NLM, for image denoising [14]. Lore et al. [17] developed LLNet, a deep auto-encoder that enhances contrast and removes noise. Zhang et al. [14] used Denoising convolutional neural networks (DnCNNs) to handle Gaussian denoising with unknown noise levels, which is different from traditional discriminative models that are trained specifically for certain noise levels. DnCNNs not only achieved excellent performance quantitatively and qualitatively by using residual learning strategy but also to speed up the training process on GPU computing by using batch normalization (BN) [18]. Zhang et al. [19] proposed a new fast, flexible CNN denoising model namely FFDNet. FFDNet can handle a wide range of noises, remove white Gaussian and spatially variant noise which requires a noise level map, and is faster than BM3D. It is effective and provides a practical solution to denoising applications because it achieves a good balance between performance and inference speed.

Although researchers have made great efforts in MRI denoising to retrieve free noise images and get effective results, the research on MRI denoising is quite limited. Current methods suffer from several drawbacks including nonlinear optimization, tuning the parameters of neural networks, high computations, and/or sensitive parameters, which seriously lead to unsatisfactory denoising results. In this work, to avoid these problems, the proposed models are inspired by an MRI denoising method based on RED-WGAN [3]. This paper mainly contributes to learning the distribution of data in a low-dimensional manifold using the WGAN framework, different loss functions such as VGG loss [20], SSL loss, SSIM loss [21], residual networks and autoencoders [22], which were employed in the proposed models to preserve clinical relevant details such as the edges and the informative structure. MSE loss in the RED-WGAN model has been replaced with SSL loss to overcome the loss of important structural details occurring due to over-smoothing edges. Also, SSIM loss has been used to preserve the image details in high resolution. The proposed method is computationally fast and can be implemented on Graphic Processing Units (GPUs). The rest of this paper is organized as follows: Section II defines the related work; Section III presents the proposed models; Section IV describes the experiments and results; and finally, Section V shows the conclusions and future work.

II. RELATED WORK

In the field of clinical imaging, Jiang et al. [23] proposed a multichannel convolutional neural network (MCDnCNN) for MRI denoising with and without a specific noise level, in which CNN layers were combined with residual learning [24] and VGG network architecture. It robustly denoises 3D MR images with Rician noise. Manjon et al. [25] proposed a two-

stage approach to effectively reduce the noise: the non-local PCA thresholding strategy is used to filter the noisy image by automatically estimating the local noise level in the image; then this filtered image is used as a guide in the rotationally invariant NLM [prefiltered rotationally invariant nonlocal means (PRI-NLM)] filter. Ran et al. [3] introduced the RED-WGAN model for MRI denoising, which consists of three main parts: the generator network, the discriminator network, and combined loss functions. In the generator network, the residual autoencoder structure is composed of convolutional and deconvolutional layers symmetrically. The discriminator network consists of convolutional layers. The authors combined three loss functions including the MSE loss function [21], Adversarial loss, and VGG loss. The proposed model powerfully reduced the noise and retrieved the structural details. Tripathi et al. [26] proposed a novel CNN-DMRI model to remove the Rician noise from MRI, which utilized a set of convolutional layers to capture the image features while the noise is separating. As part of CNN-DMRI structures, encoder-decoder structures were also employed to retain the informative features of the image while unnecessary ones are ignored. The qualitative and quantitative results of the proposed method are promising. Li et al. [1] successfully applied Rician denoising with a progressive learning approach to MR images. The progressive network, called RicianNet, consists of two sub-RicianNets, which are residual blocks: one of the sub-networks fitted the noise distribution at the pixel-domain without batch normalization layer, and the other one employs ResNet structure with batch normalization layer in the feature domain, thus enhancing the nonlinear mapping. The authors improved the network performance by employing the BN layer, Convolutional layer, and residual unit. RicianNet had better quantitative measures and significant improvements in visual inspections. Aetesam et al. [27] proposed a deep neural architecture for MRI denoising to remove Gaussian-impulse noise by using an ensemble-based residual learning strategy. The proposed model achieved high-quality visual results and high quantitative metrics compared to other state-of-the-art models. Gregory et al. [28] developed a multi-branch deep neural network architecture, HydraNet, to remove noise from MR images at a wide range of noise levels. Compared to other deep learning-based methods, the HydraNet network demonstrated powerful results in the denoising of complex noise distributions. Wu et al. [2] used 3D Parallel-RicianNet for 3D MRI denoising, which combines global and local information for noise reduction. To expand its receptive field, the authors introduced a powerful module called dilated convolution residuals (DCR).

III. PROPOSED DENOISING MODEL

It is difficult to denoise an MRI because magnitude images, which consist of real and imaginary parts, are commonly used [3, 29]. The noise in the magnitude MR image follows a Rician noise distribution [29], which is significantly more complicated than traditional additive Gaussian noise.

In MRI denoising, a free MR image is obtained by removing noise from a noisy MR image, as follows:

$$z = \theta(x) \quad (1)$$

Where z denotes a noisy MR image, x denotes the corresponding noise-free MR image and $(x, z \in R^{H \times W \times D})$, and function θ maps to the noise. The model-based DL is independent of noise and its statistical characteristics since it is a black box. So, the denoising process aims to approximate the function θ^{-1} to the possible optimal and can be expressed as follows:

$$\operatorname{argmin}_f \|\hat{z} - z\|_2^2 \quad (2)$$

where $\hat{z} = Q(z)$, which corresponds to an estimate of x , and Q indicates the optimal approximation of θ^{-1} [3].

A. Wasserstein GAN

The GANs model can be described in Eq. (3).

$$\min_G \max_D V(D, G) = E_{x \sim P_r(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

The two variables x and z can be interpreted as samples drawn from two different distributions of data, which in the context of statistics, can be considered as being real image distributions p_r and noisy image distributions p_z respectively. Then, the denoising function moves the samples from p_z to p_g which is close to p_r . An Adversarial Generative Network (GAN) is made up of two networks, a generator, G , and a discriminator, D . There have been many uses of GAN in research fields such as computer vision [30,31], security [32], and data generation [33]. The generator generates samples from random noise as close as possible to real data to deceive a discriminator. The discriminator attempts to distinguish between the two-distribution of the generative model p_g and the real data p_r .

Despite its success in image generation, GAN suffers from training instabilities, extremely sensitive parameter tuning, vanishing gradient, and mode collapse [34]. It has been proposed to improve GAN by using Wasserstein GAN (WGAN) [35]. The loss function of WGAN was proposed to avoid vanishing gradients. Wasserstein Distance measures the divergence between real distribution P_r and model distribution P_g ; In WGAN, weight clipping is used to enforce Lipschitz constraints, when clipping parameters are too small or too large can result in the same original GAN problems. Therefore, the Gradient penalty (GP) was used instead of weight clipping to enforce the Lipschitz constraint on the critic(discriminator) during training.

WGAN-GP is a WGAN with a gradient penalty, and the loss function is shown in Eq. (4).

$$L = E_{\tilde{x} \sim p_g} [D(\tilde{x})] - E_{x \sim p_r} [D(x)] + \lambda E_{\hat{x} \sim p_{\hat{x}}} [(\|\nabla_{\tilde{x}} D(\hat{x})\|_2 - 1)^2] \quad (4)$$

The sample of $p_{\hat{x}}$ is taken uniformly between two points sampled from P_r and P_g , the last term is a gradient penalty factor, and λ is a penalty coefficient.

B. Loss Functions

Mean Squared Error (MSE) loss or $L2$ calculates the normalized Euclidean distance between a generated patch $G(z)$ from model distribution P_g and the patch of noise-free images x from real distribution P_r ; it minimizes the pixel-wise difference

between them [21]. Recent studies suggest that although the per-pixel MSE results have a high peak signal-to-noise ratio (PSNR), it may cause the loss of some important structural details due to an over-smoothed edge. The formula of $L2$ loss is expressed as in Eq. (5).

$$L2 = \frac{1}{HWD} \|G(z) - x\|_2^2 \quad (5)$$

The Perceptual Loss (PL) was used to overcome this problem by being employed in the feature space instead of directly estimating MSE on a pixel-by-pixel basis. A pre-trained VGG-19 network [20] can be applied to extract the features from the generated patch and noise-free patch, VGG loss compares high-level perceptual differences [21].

$$PL(G) = E_{(x,z)} \frac{1}{HWD} \|\varphi G(z) - \varphi(x)\|_2^2 \quad (6)$$

In which φ is a feature extractor, W refers to the width, H indicates to the height, and D is the depth of feature maps. The perceptual loss can be described as the following formula:

$$L_{VGG}(G) = PL(G) E_{(x,z)} \frac{1}{HWD} \|VGG(G(z)) - VGG(x)\|_2^2 \quad (7)$$

Structural Similarity (SSIM) loss measures the similarity between two patches $G(z)$ and x based on three comparisons: contrast, luminance, and structure [21]. The SSIM can perform better than the MSE in perceptual pattern recognition because it is visually based. The original SSIM is formulated as follows in Eq. (8).

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C1}{\mu_x^2 + \mu_y^2 + C1} * \frac{2\sigma_{xy} + C2}{\sigma_x^2 + \sigma_y^2 + C2} = L(x, y) * cs(x, y) \quad (8)$$

Where μ_x , μ_y , σ_x , σ_y and σ_{xy} are the means, standard deviations, and the cross-covariance of the two images (y, x) obtained from the model and the corresponding noise-free image respectively and $C1$, $C2$ are constants [21,36,37]. If x and y are very similar, SSIM approaches 1.

$$L_{SSIM} = 1 - SSIM(x, y) \quad (9)$$

In this paper, we presented RED-WGAN-SSL and RED-WGAN-SSIM models based on WGAN. They are incorporated with different loss functions to reduce the noise in 3D MRI and retain structural information. The two proposed models are compared with RED-WGAN [3]. The joint loss functions for all models are formulated as follows:

$$L_{RED-WGAN} = \lambda_1 L_{WGAN}(G) + \lambda_2 L_{VGG} + \lambda_3 L_{MSE} \quad (10)$$

$$L_{RED-WGAN-SSIM} = \lambda_1 L_{WGAN}(G) + \lambda_2 L_{VGG} + \lambda_3 L_{SSIM} \quad (11)$$

$$L_{RED-WGAN-SSL} = \lambda_1 L_{WGAN}(G) + \lambda_2 L_{VGG} + \lambda_3 L_{SSL} \quad (12)$$

C. Network Architectures

The proposed models' architecture is inspired by the RED-WGAN architecture [3], which is made up of a G network, a D network, and a VGG network. The G network structure is an autoencoder network that consists of the convolution and deconvolution layers that are symmetrical to deal with the noise. The convolution and deconvolution layer pairs are linked by short connections. The deconvolution layers and the short connections are proposed to speed up the training procedure and maintain more details. There are 8 layers in the

encoder-decoder generator: four convolutional layers and four deconvolutional layers. A 3D convolution is applied to the first seven layers, followed by a batch-normalization and a LeakyReLU, except the last layer, which has a 3D convolution and a LeakyReLU without a batch-normalization; each layer uses $3 \times 3 \times 3$ kernels, the generator employed 32, 64, 128, 256, 128, 64, 32, 1 filter. The VGG network is used to extract features.

The structure of the discriminator network D consists of 3 convolutional layers. All layers perform 3D convolutional operations in sequence with 32, 64 and 128 filters and have $3 \times 3 \times 3$ kernel size, followed by a fully connected layer in the last layer that has a single output.

IV. EXPERIMENTS AND RESULTS

The two proposed models RED-WGAN-SSL and RED-WGAN-SSIM were extensively tested on clinical datasets to validate their performance.

A. Clinical Data

Clinical experiments were conducted using the IXI dataset [38] gathered from three hospitals: Hammersmith Hospital, Guy's Hospital, and the Institute of Psychiatry. The above-mentioned website provides detailed information on scanning configuration. The Hammersmith dataset is a subset of the IXI dataset obtained from a Philips 3T scanner. 110 PD-weighted brain image volumes were randomly chosen. The training set consists of 100 image volumes from the Hammersmith dataset, and the testing set consists of 5 image volumes from the Hammersmith dataset, it also included 5 image volumes from the Guy's Hospital dataset to evaluate the robustness of the proposed models. We manually added Rician noise to the training set and testing set to simulate noisy images. Many training samples are required for deep learning-based methods, which is especially challenging in clinics.

B. Parameter Setting

The training was performed on PD-weighted brain image volumes with specific levels of noise. According to the suggestions in [31,39], the parameters λ_1 , λ_2 , and λ_3 were experimentally set to 1, 0.1, and $1e-3$, respectively. A penalty coefficient λ in Eq. (4) was specified in following the suggestion [35] to 10. The loss function was optimized by the Adam algorithm [40], and the parameters α , β_1 , and β_2 for the optimizer were set to $1e-4$, 0.5, and 0.9, respectively.

C. Results

To evaluate the performance of the proposed denoising models RED-WGAN-SSL and RED-WGAN-SSIM in comparison to RED-WGAN, three quantitative metrics were utilized. The first metric, PSNR, involved comparing the denoised images to the original (ground truth) images by calculating RMSE. The second metric, RMSE, measured the difference between the denoised and ground truth images, lower values indicating better image quality. Lastly, the SSIM was used to compare the similarities between the denoised and ground truth images, taking into account the luminance, contrast, and structure of the images.

1) *Results obtained using a mini-batch size of 11*: This section illustrates the different results for RED-WGAN-SSL, RED-WGAN-SSIM, and RED-WGAN that were trained on PDw images with different levels of noise (5%, 9%, 11%, and 15%). Then, the three denoising models were tested on the same levels of noise (5%, 9%, 11%, and 15%).

a) *Quantitative Results*: Table I presents the average quantitative analysis. The results demonstrate that when the noise level is less than 11%, RED-WGAN-SSL and RED-WGAN-SSIM exhibit slightly better performance than RED-WGAN. As the noise level increases, the performance of RED-WGAN-SSL is mildly better than that of RED-WGAN and RED-WGAN-SSIM.

b) *Qualitative Results*: This section illustrates the different qualitative results for the denoising models RED-WGAN-SSL, RED-WGAN-SSIM and RED-WGAN. Fig. 1 shows results obtained for the PDw brain images in the testing set with 15% Rician noise as the models were also trained on images with 15% Rician noise. Each model suppresses noise to a different degree. However, some vital details are distorted as in RED-WGAN-SSIM. In Fig. 2, it is important to mention that all the models at the noise level of 11% can remove noise to a different degree and that the RED-WGAN-SSL and RED-WGAN-SSIM have better results compared with RED-WGAN, as they preserve more structural details than RED-WGAN as shown by the red arrow. RED-WGAN-SSL suppresses noise better than other models. The results show that the lower the noise level, the better the results and closer to the original reference image as observed at level noise of 9% and 5% in Fig. 3 and Fig. 4 respectively. Consequently, the structure details were preserved while noise was effectively reduced especially at level noise of 5%.

2) *Results obtained using a mini-batch size of 80*: Based on the results obtained in Table II, the RED-WGAN-SSL seems to have performed better in terms of PSNR, SSIM, and RMSE than all the models considered. Fig. 5 provides a visual representation of the different results for RED-WGAN-SSL, RED-WGAN-SSIM, and RED-WGAN on the PDw brain images that were corrupted by 15% Rician noise in the training set and then were tested with 15% Rician noise. It is important to note that all of the models are capable of suppressing noise in converging degrees. The RED-WGAN-SSL model has an improvement in noise suppression compared to the RED-WGAN model as shown in the red arrow, and it also produces results that are more consistent compared to the original reference images. RED-WGAN-SSL analysis results show that most of the noise has been reduced efficiently and the structure details have been retained much better than other models. The quantitative results of different models for Fig. 5 are presented in Table II. There was an agreement between the visual inspection and quantitative results in terms of PSNR, SSIM, and RMSE when using RED-WGAN-SSL, which is the best result of all the modalities.

TABLE I. A COMPARISON OF PSNR, SSIM, AND RMSE METRICS ON DENOISED PDW EXAMPLE WITH DIFFERENT LEVELS OF RICIAN NOISE FROM THE TESTING SET

		5%	9%	11%	15%
Noise	PSNR	41.853721	46.758666	44.828463	41.853721
	SSIM	0.306874	0.192677	0.157860	0.109305
	RMSE	0.374527	0.678883	0.831107	1.135392
RED-WGAN	PSNR	63.428257	59.420627	57.281444	54.901224
	SSIM	0.802056	0.760506	0.699665	0.605663
	RMSE	0.116628	0.180583	0.229623	0.303276
RED-WGAN-SSL	PSNR	63.783810	59.621328	57.85996	55.001128
	SSIM	0.816764	0.761466	0.721758	0.588793
	RMSE	0.112812	0.177403	0.215451	0.301685
RED-WGAN-SSIM	PSNR	63.617620	59.364931	57.574540	54.617207
	SSIM	0.787149	0.726462	0.718152	0.560422
	RMSE	0.114135	0.181923	0.222884	0.313424

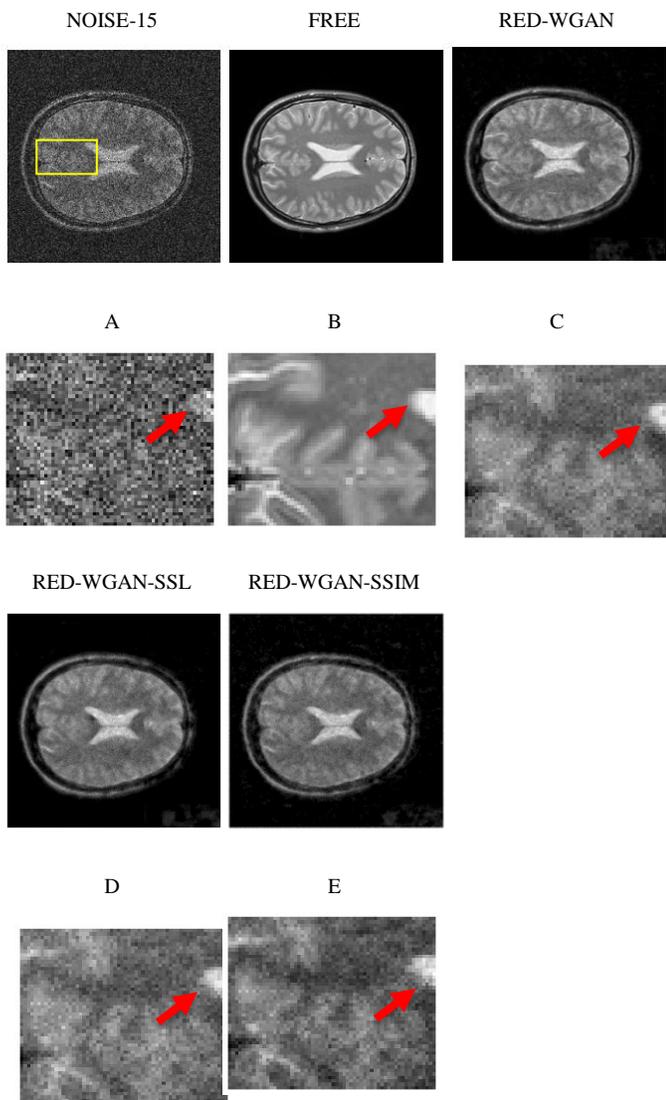


Fig. 1. Denoised PDw example with 15% Rician noise from the testing set at a mini-batch size =110: (A) Noisy image, (B) Ground truth image, (C) RED-WGAN, (D) RED-WGAN-SSL and (E) RED-WGAN-SSIM.

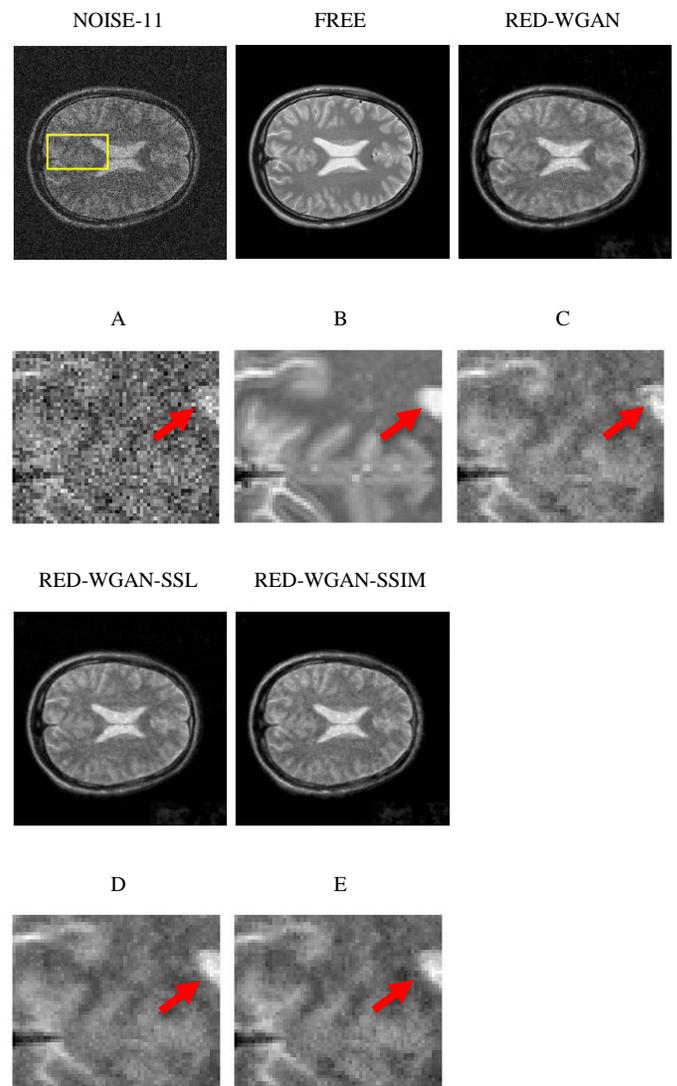


Fig. 2. Denoised PDw example with 11% Rician noise from the testing set at a mini-batch size =110: (A) Noisy image, (B) Ground truth image, (C) RED-WGAN, (D) RED-WGAN-SSL and (E) RED-WGAN-SSIM.

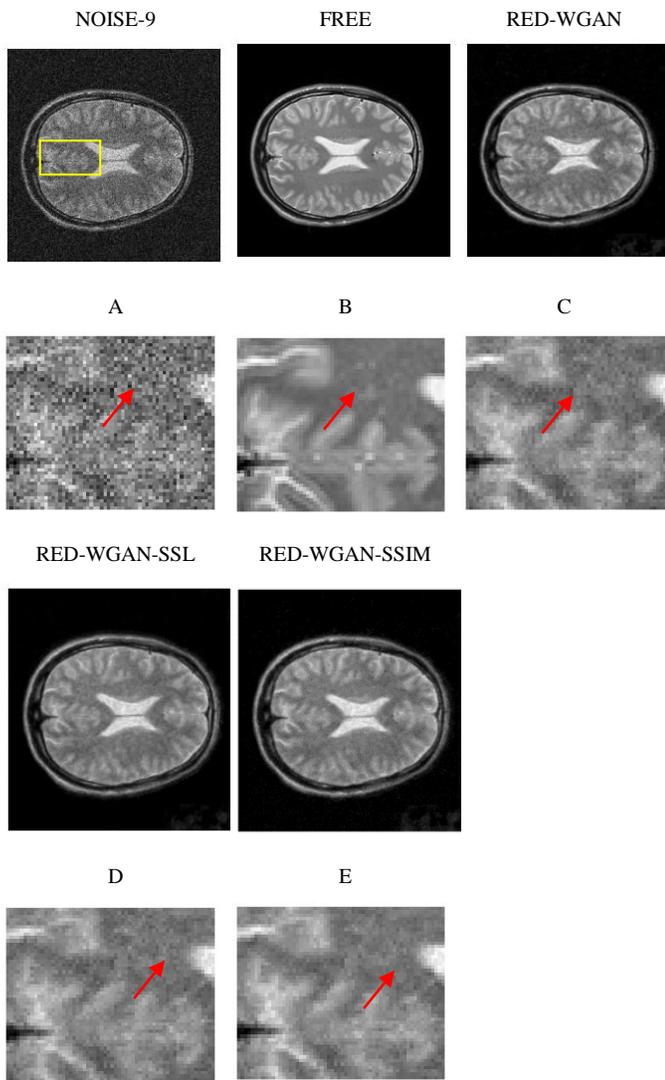


Fig. 3. Denoised PDw example with 9% Rician noise from the testing set at a mini-batch size =110: (A) Noisy image, (B) Ground truth image, (C) RED-WGAN, (D) RED-WGAN-SSL and (E) RED-WGAN-SSIM.

TABLE II. A COMPARISON OF PSNR, SSIM AND RMSE METRICS ON DENOISED PDW EXAMPLE WITH 15% RICIAN NOISE FROM THE TESTING SET AT A MINI-BATCH SIZE =80

		PSNR	41.853721
	Noise -15	SSIM	0.109305
		RMSE	1.135392
	RED-WGAN	PSNR	58.497804
		SSIM	0.755688
		RMSE	0.212351
	RED-WGAN-SSL	PSNR	58.673452
		SSIM	0.781521
		RMSE	0.205079

		PSNR	58.342468
	RED-WGAN-SSIM	SSIM	0.736552
		RMSE	0.212479

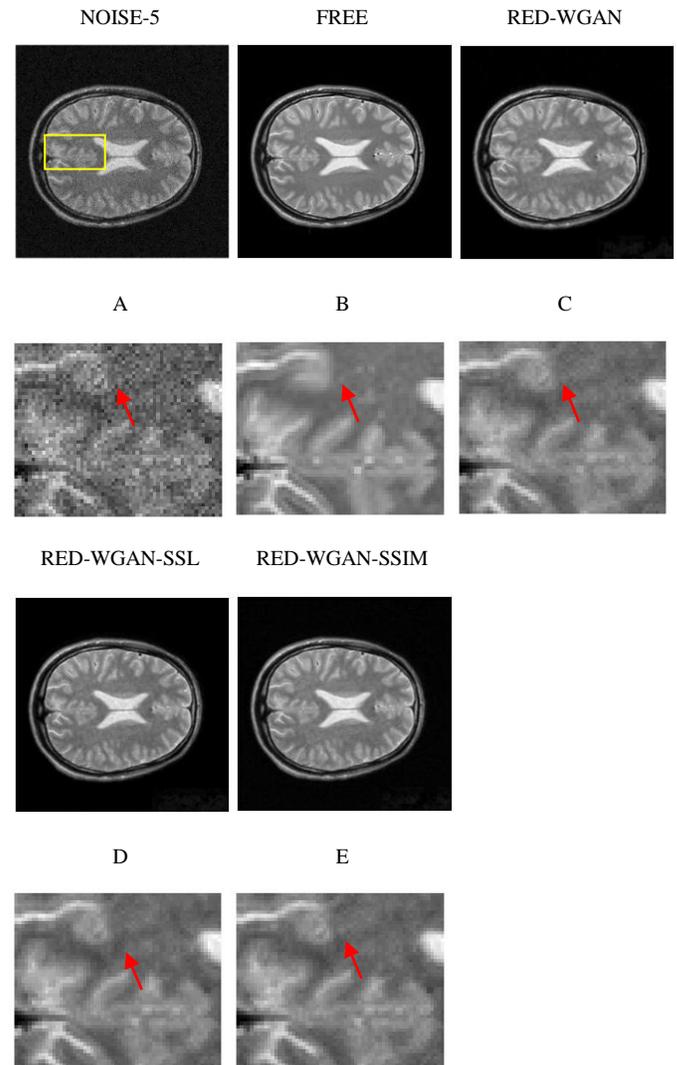


Fig. 4. Denoised PDw example with 5% Rician noise from the testing set at a mini-batch size =110: (A) Noisy image, (B) Ground truth image, (C) RED-WGAN, (D) RED-WGAN-SSL and (E) RED-WGAN-SSIM.

3) Comparison between the results obtained with a mini-batch size = 80 and mini-batch size = 110: Quantitative results of all models at a noise level of 15% with a mini-batch size of 80 are significantly better than those with a mini-batch size of 110 as shown in Table I and Table II. Based on the comparison, we found that the qualitative results with a mini-batch of 80 show that most noise can be effectively removed in most cases, as well as that the structural details are preserved better than the results with a mini-batch of 110 as shown in Fig. 5

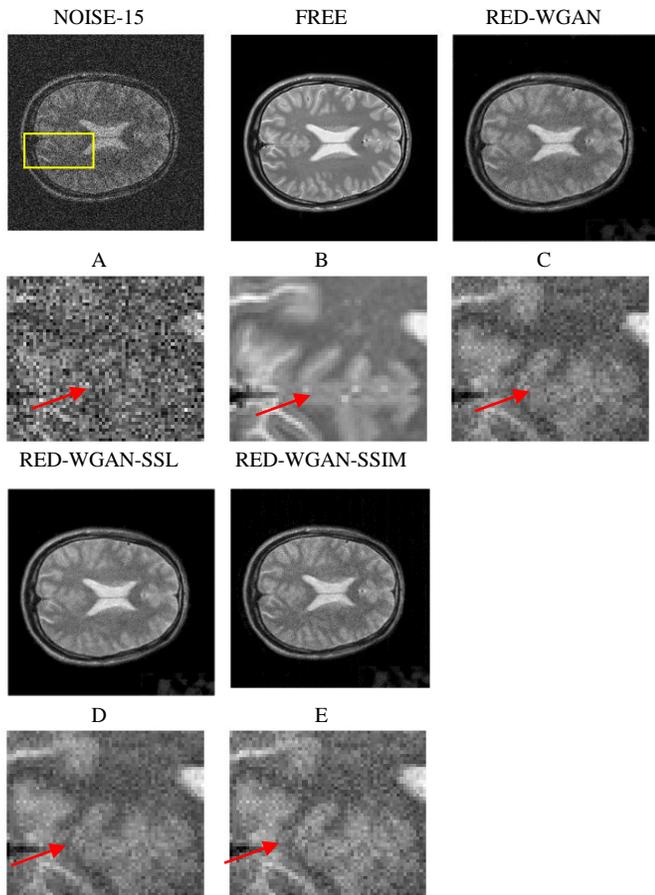


Fig. 5. Denoised PDw example with 15% Rician noise from the testing set at a mini-batch size =80: (A) Noisy image, (B) Ground truth image, (C) RED-WGAN, (D) RED-WGAN-SSL and (E) RED-WGAN-SSIM.

4) *An evaluation of robustness:* For the analysis of the robustness of the RED-WGAN-SSL and RED-WGAN-SSIM

models for various noise levels, RED-WGAN-SSL, RED-WGAN-SSIM and RED-WGAN models were trained with 15% Rician noise, and these three models were then tested with various noise levels which are 5%, 9%, 11%, 15%, and 17% to show how robust they are.

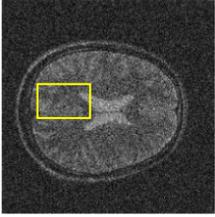
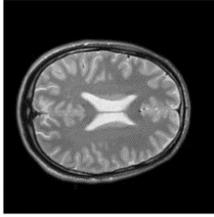
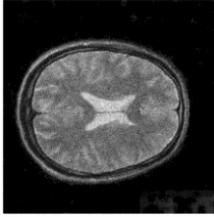
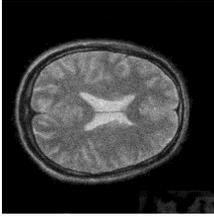
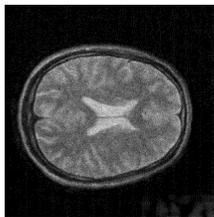
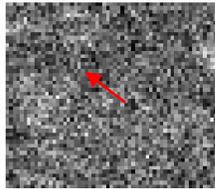
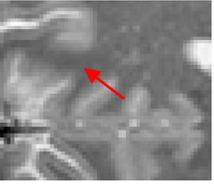
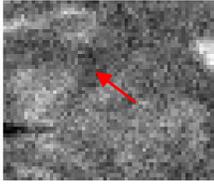
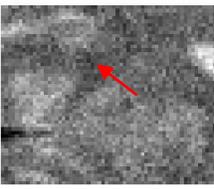
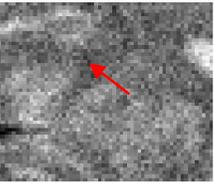
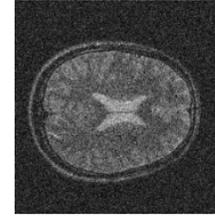
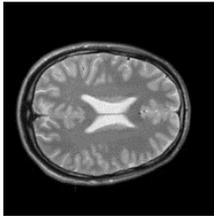
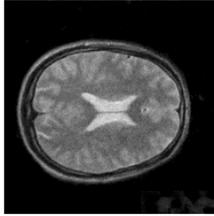
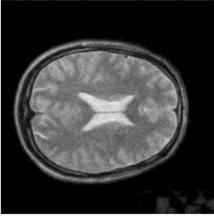
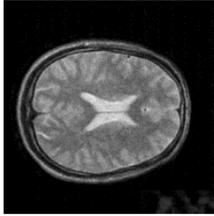
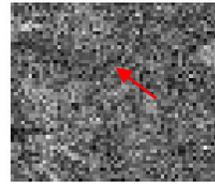
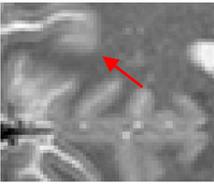
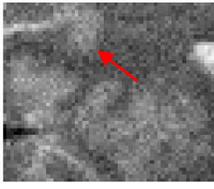
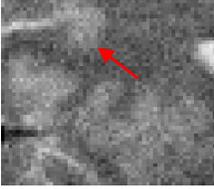
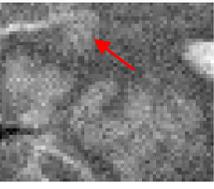
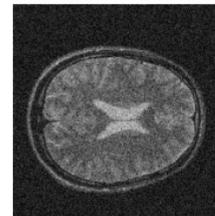
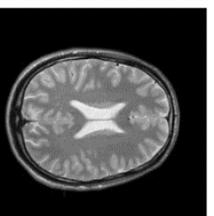
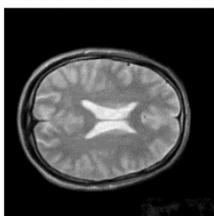
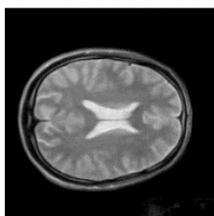
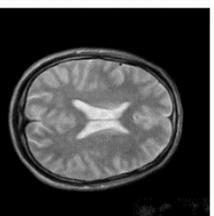
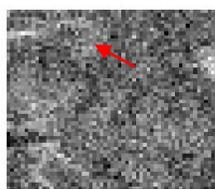
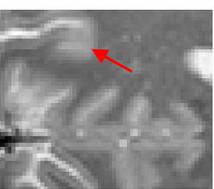
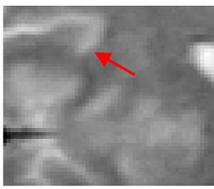
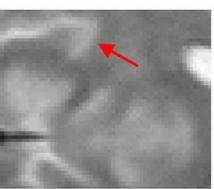
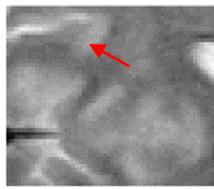
a) *Qualitative Results:* It has been found that the performance of RED-WGAN-SSL is slightly better than that of other models at a higher level than 11% Rician noise. RED-WGAN-SSIM preserved more informative features and provided better visual quality compared to other denoising on the testing set at a lower level than 11% Rician noise, it can reduce the noise and artifacts as indicated by the red arrow in Fig. 6.

The advantage of our proposed models is combining loss functions which are SSL, SSIM, and VGG losses. The VGG loss is used to preserve image style and content after it has been denoised. SSL loss is efficient in extracting structural details and informative features. The SSIM loss generates visually artistic images by using the visible structures in the image. All of these losses help to generate results that can be similar to the original distribution of the data.

b) *Quantitative Results:* A quantitative summary of the results for Fig. 6 is provided in Table III. The RED-WGAN-SSL has better scores when tested at noise levels higher than 11%; it provides good PSNR and SSIM values, which are higher than those of other models as observed in Table III. The RED-WGAN-SSIM has better scores when tested at noise levels less than 11%. As a result, we can take this as evidence that the proposed models are both robust and generalizable. Consequently, we can conclude that the proposed models can denoise MR images with high-quality images and with high structural similarity between the original image and its denoised result.

TABLE III. A COMPARISON OF PSNR, SSIM, AND RMSE MEASURES ON PDW IMAGES WITH DIFFERENT NOISE LEVELS IS SHOWN FROM TOP TO BOTTOM

Noise	RED-WGAN-15	RED-WGAN -SSIM-15	RED-WGAN -SSL-15	Noise
5%	52.446672	53.204429	53.375807	52.93542
	0.306874	0.664977	0.684727	0.665400
	0.374527	0.350747	0.343258	0.361404
9%	46.758666	56.432396	56.666630	56.06631
	0.192677	0.734335	0.745195	0.733152
	0.678883	0.254213	0.247617	0.264126
11%	44.828463	58.067011	58.161415	57.496115
	0.157860	0.763594	0.775136	0.762014
	0.831107	0.214020	0.210872	0.228062
15%	41.853721	58.497804	58.342468	58.673452
	0.109305	0.755688	0.736552	0.781521
	1.135392	0.212351	0.212479	0.205079
17%	40.643732	57.271908	56.900920	57.800380
	0.091793	0.713646	0.644451	0.758627
	1.289173	0.252457	0.257386	0.231039
19%	39.576729	55.302866	54.905874	56.213300
	0.077927	0.626172	0.539672	0.713555
	1.442252	0.315957	0.326715	0.282167

Noise Level	NOISY	FREE	RED-WGAN-15	RED-WGAN-SSL-15	RED-WGAN-SSIM-15
17%					
					
15%					
					
11%					
					

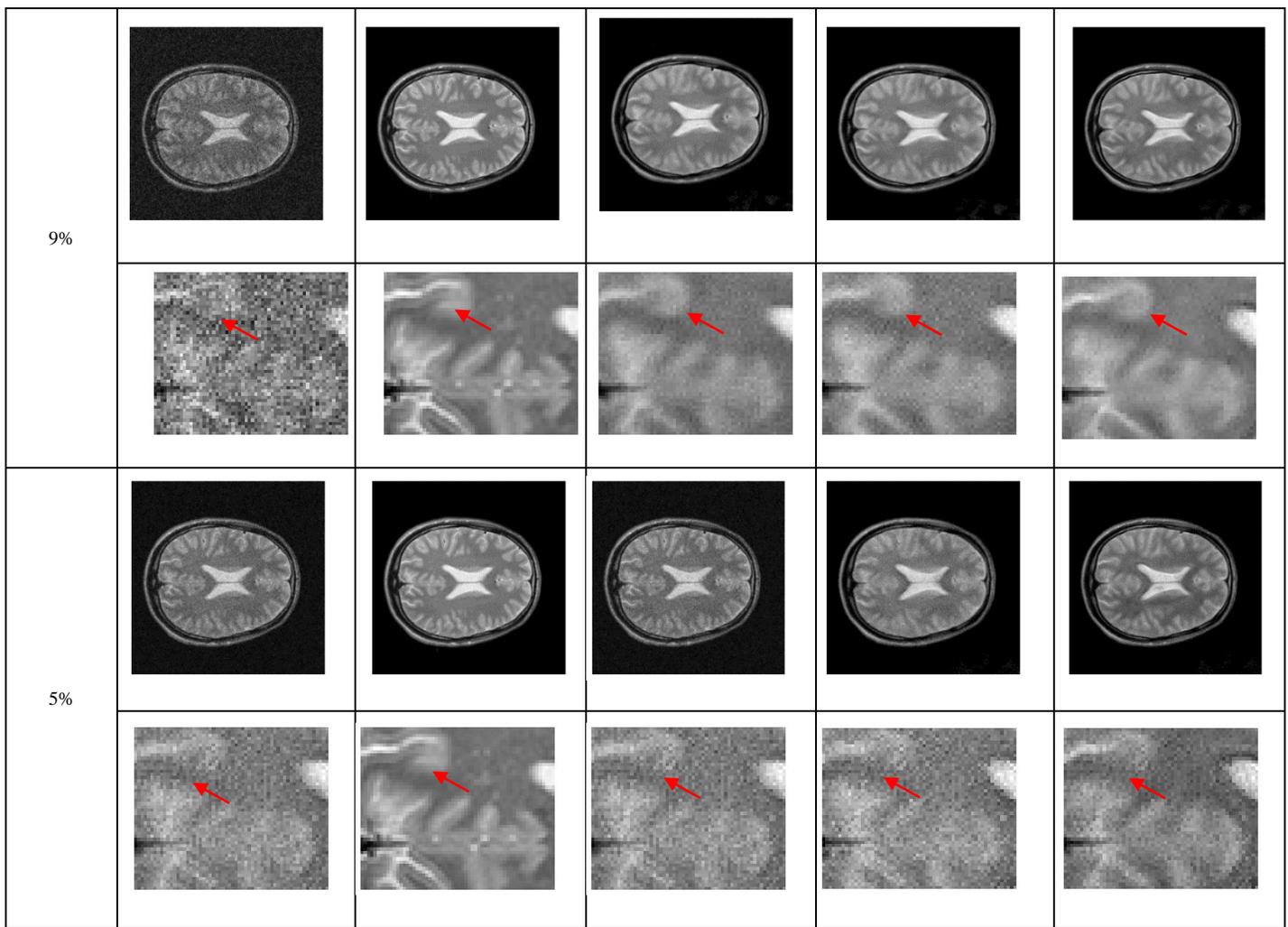


Fig. 6. Denoised PDw example with different levels of Rician noise from the testing set. (A) Noisy image, (B) Ground truth image, (C) RED-WGAN-15, (D) RED-WGAN-SSL-15 and (E) RED-WGAN-SSIM-15.

V. CONCLUSION

The two models RED-WGAN-SSL and RED-WGAN-SSIM models were presented in this paper, which use WGAN to get rip Rician noise from MR images while maintaining structure details. A 3D CNN has been used in these models to process 3D volume data. As well as using the WGAN framework, we introduced an autoencoder generator structure and combined loss functions. We have also improved the performance of our models by adapting the mixture of SSL, SSIM, and VGG loss functions. According to the results of the experiments, the performance of RED-WGAN-SSL and RED-WGAN-SSIM, which are based on the WGAN, SSL, SSIM, and perceptual loss, have been significantly improved both qualitatively and quantitatively. Compared with the RED-WGAN model, they can suppress the noise at the same time as retaining a higher level of detail. A comparison of the results of all models at a noise level of 15% when a mini-batch size = 80 is superior to a mini-batch size = 110. It is interesting to note that the RED-WGAN-SSL scores better metrics on the testing set at noise levels higher than 11%, whereas the RED-WGAN-SSIM scores better metrics on the testing set at noise levels less

than 11%. This leads us to conclude that our proposed models are both robust and generalizable and can therefore be viewed as a strong indication that our work is well worth the effort. A deep learning-based method has a high computational cost. Most of the costs are incurred during the training stage. Although most training is conducted on a GPU, it still takes a long time. In future work, the proposed models will be implemented in T1 and T2 brain image volumes. As well as this, we will apply our denoising methods to a variety of medical images with different types of noise.

REFERENCES

- [1] S. Li, J. Zhou, D. Liang and Q. Liu, "MRI denoising using progressively distribution-based neural network, *Magnetic resonance imaging*, vol. 71, pp. 55-68, 2020.
- [2] L. Wu, S. Hu and C. Liu, "Denoising of 3D Brain MR Images with Parallel Residual Learning of Convolutional Neural Network Using Global and Local Feature Extraction," *Computational Intelligence and Neuroscience*, pp. 1-18, 2021.
- [3] M. Ran, J. Hu, Y. Chen, H. Chen, H. Sun, et al., "Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network. *Medical image analysis*, vol. 55, pp. 165-180, 2019.

- [4] J. Yang, J. Fan, D. Ai, S. Zhou, S. Tang et al., "Brain MR image denoising for Rician noise using pre-smooth non-local means filter. Biomedical engineering online, vol. 14, no. 1, pp. 1-20, 2015.
- [5] E. Arias Castro and D.L. Donoho "Does median filtering truly preserve edges better than linear filtering?" vol. 37, no. 3, pp. 1172-1206, 2009.
- [6] R.A. Haddad and A.N. Akansu, "A class of fast Gaussian binomial filters for speech and image processing," IEEE Transactions on Signal Processing, vol. 39, no. 3, pp.723-727, 1991.
- [7] J. Chen, J. Benesty, Y. Huang and S. Doclo "New insights into the noise reduction Wiener filter," IEEE Transactions on Audio, Speech and Language Processing, vol. 14, no. 4, pp. 1218-1234, 2006.
- [8] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic Diffusion," IEEE Trans. Pattern Anal. Mach. Intell, vol. 12, no. 7 pp. 629-63, 1990.
- [9] G. Gerig, O. Kubler, R. Kikinis and F.A. Jolesz, "Nonlinear anisotropic filtering of MRI data," IEEE Trans. Med. Imaging, vol. 11, no. 2 pp. 221-232, 1992.
- [10] J. E. Fowler, "The redundant discrete wavelet transform and additive noise," IEEE Signal Processing Letters, vol. 12, no. 9, pp. 629-632, 2005.
- [11] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Transactions on image processing, vol. 16, no. 8, pp. 2080-2095, 2007.
- [12] X. Zhang, Z. Xu, N. Jia, W. Yang, Q. Feng et al., "Denoising of 3D magnetic resonance images by using higher-order singular value decomposition," Medical image analysis, vol. 19, no. 1, pp.75-86, 2015.
- [13] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P. A. Manzagol, et al., "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," Journal of machine learning research, vol. 11, no. 12, pp. 1-38, 2010.
- [14] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," IEEE transactions on image processing, vol. 26, no. 7, pp. 3142-3155, 2017.
- [15] X. Tao, H. Gao, X. Shen, J. Wang and J. Jia, "Scale-recurrent network for deep image deblurring," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8174-8182, 2018.
- [16] C. Dong, C.L. Chen, K. He, X. Tang, "Image super-resolution using deep convolutional networks," IEEE Trans. Pattern Anal. Mach. Intell, vol. 38, no. 2, pp. 295-307, 2015.
- [17] K. G. Lore, A. Akintayo and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," Pattern Recognition, vol. 61, no. 3, pp. 650-662, 2017.
- [18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," In International conference on machine learning, vol. 37, pp. 448-456, 2015.
- [19] K. Zhang, W. Zuo and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," IEEE Transactions on Image Processing, vol. 27, no. 9, pp. 4608-4622, 2018.
- [20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, pp.1-14, 2014, [Online]. Available: <https://arxiv.org/abs/1409.1556>.
- [21] H. Zhao, O. Gallo, I. Frosio and J. Kautz, "Loss functions for image restoration with neural networks," IEEE Transactions on Computational Imaging, vol. 3, no. 1, pp. 47-57, 2016.
- [22] D. Bank, N. Koenigstein and R. Giryes, "Autoencoders," arXiv preprint arXiv:05991, pp. 1-22, 2021.
- [23] D. Jiang, W. Dou, L. Vosters, X. Xu, Y. Sun et al., "Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network," Japanese journal of radiology, vol. 36, pp. 566-574 (2018).
- [24] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
- [25] J.V. Manjón, P. Coupé and A. Buades, "MRI noise estimation and denoising using non-local PCA," Medical image analysis, vol. 22, no. 1, pp. 35-47, 2015.
- [26] P. C. Tripathi and S. Bag, "CNN-DMRI: a convolutional neural network for denoising of magnetic resonance images," Pattern Recognition Letters, vol. 135, no. 1, pp. 57-63, 2020.
- [27] H. Aetesam and S. K. Maji, "Noise-dependent training for deep parallel ensemble denoising in magnetic resonance images," Biomedical Signal Processing and Control, vol. 66, pp. 102405-102415, 2021.
- [28] S. Gregory, H. Cheng, S. Newman and Y. Gan, "HydraNet: a multi-branch convolutional neural network architecture for MRI denoising," In Medical Imaging 2021: Image Processing, vol. 11596, pp. 881-889, 2021.
- [29] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," Magnetic Resonance. Med. vol. 34, no. 6, pp. 910 - 914,1995.
- [30] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125-1134. 2017.
- [31] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA, pp. 4681-4690. 2017.
- [32] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure Steganography Based on Generative Adversarial Networks," In Pacific Rim Conference on Multimedia, pp. 534-544, 2017.
- [33] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," arXiv preprint arXiv:1511.06434, 2015, [Online]. Available: <https://arxiv.org/abs/1511.06434>.
- [34] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford et al., "Improved techniques for training GANs," Advances in Neural Information Processing Systems, Barcelona, Spain, vol. 29, pp. 1-9, 2016.
- [35] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumouli and A. Courville, "Improved training of Wasserstein GANs," Advances in Neural Information Processing Systems, Long Beach, CA, USA, vol. 30, pp. 5768-5778, 2017.
- [36] C. You, Q. Yang, H. Shan, L. Gjestebj, G. Li et al., "Structurally-sensitive multi-scale deep neural network for low-dose CT denoising," IEEE Access, vol. 6, pp. 41839-41855, 2018.
- [37] A. A. Mahmoud, H. A. Sayed and S. S. Mohamed, "Variant wasserstein generative adversarial network applied on low dose ct image denoising," Computers, Materials & Continua, vol. 75, no.2, pp. 4535-4552, 2023.
- [38] <http://brain-development.org/ixi-dataset/>, Accessed 10 March. 2022.
- [39] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi et al., "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1348-1357, 2018.
- [40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, pp.1-15, 2014, [Online]. Available: <https://arxiv.org/abs/1412.6980>.