



NOISE REMOVAL IN MAGNETIC RESONANCE IMAGING USING 3D DEEP LEARNING MODEL

Rukiye KARAKIS*, Faculty of Technology, Software Engineering Department, Sivas Cumhuriyet University, Sivas/Turkey, DEEPBRAIN: Neuro-Imaging and Artificial Intelligence Research Group, Sivas Cumhuriyet University, Sivas, Turkey, rkarakis@cumhuriyet.edu.tr

 <https://orcid.org/0000-0002-1797-3461>

Tugba TOPDAG, Department of Computer Engineering, Institute of Science, Sivas Cumhuriyet University, Sivas/Turkey, tugbatopdag@hotmail.com

 <https://orcid.org/0009-0003-7697-1970>

Received: 06.08.2024, Accepted: 25.10.2024

Research Article

*Corresponding author

DOI: 10.22531/muglajsci.1527803

Abstract

Magnetic Resonance Imaging (MRI) is a widely used imaging technique for examining brain tissues and diagnosing various conditions. However, MRI images often contain noise caused by factors such as equipment limitations, environmental conditions, patient movement, and magnetic field interference. This noise can obscure critical details, making accurate diagnosis and treatment planning challenging. In this study, the focus is on the removal of Rician noise from MRI images. To address this challenge, two 3D autoencoder models, named M-UNet+ResNet and M-UNet+DenseNet, were developed. These models are based on an enhanced UNet architecture that integrates dense and residual connections, aimed at improving noise reduction capabilities. The models were trained using T1 and T2-weighted MRI images from the IXI dataset, incorporating noise levels varying from 3% to 15%. Their performance was evaluated using metrics such as peak signal-to-noise ratio, structural similarity index measure, and mean absolute error. The results demonstrated that both models effectively reduced noise across various levels, with M-UNet+ResNet generally outperforming M-UNet+DenseNet. Notably, M-UNet+ResNet achieved PSNR values of 38.72 dB and 37.04 dB, and SSIM values of 0.82 and 0.81 in the IXI-HH-T2 and IXI-Guys-T2 datasets, respectively, indicating its strong capability in preserving image quality. This study concludes that incorporating residual connections in DL models enhances their ability to remove noise from MRI images, offering a solution for maintaining the integrity of medical images in clinical settings.

Keywords: Noise removal, Magnetic resonance imaging, Deep learning, Residual connection, UNet

MANYETİK REZONANS GÖRÜNTÜLEMEDE GÜRÜLTÜ GİDERME İÇİN 3D DERİN ÖĞRENME MODELİ

Özet

Manyetik Rezonans Görüntüleme (MRI), beyin dokularını incelemek ve çeşitli durumları teşhis etmek için yaygın olarak kullanılan bir görüntüleme tekniğidir. Ancak, MRI görüntüleri genellikle cihaz kısıtlamaları, çevre koşulları, hasta hareketi ve manyetik alan girişimi gibi faktörlerin neden olduğu gürültüleri içerir. Bu gürültü kritik ayrıntıları gizleyebilir ve doğru tanı ve tedavi planlamasını zorlaştırabilir. Bu çalışmada, MRI görüntülerinden Rician gürültüsünün giderilmesine odaklanılmıştır. Bu zorluğun üstesinden gelmek için M-UNet+ResNet ve M-UNet+DenseNet adlı iki 3B otokodlayıcı modeli geliştirilmiştir. Bu modeller, gürültü azaltma yeteneklerini iyileştirmeyi amaçlayan yoğun ve kalıntı bağlantıları entegre edilerek geliştirilmiş bir UNet mimarisine dayanmaktadır. Modeller, %3 ila %15 arasında değişen gürültü seviyelerine sahip IXI veri setinden T1 ve T2 ağırlıklı MRI görüntüleri üzerinde eğitilmiştir. Modellerin performansları, tepe sinyal-gürültü oranı, yapısal benzerlik indeksi ölçümü ve ortalama mutlak hata gibi ölçütler kullanılarak değerlendirilmiştir. Sonuçlar, her iki modelin de çeşitli seviyelerde gürültüyü etkili bir şekilde azalttığını ve M-UNet+ResNet'in genel olarak M-UNet+DenseNet'ten daha iyi performans gösterdiğini göstermiştir. Özellikle, M-UNet+ResNet, IXI-HH-T2 ve IXI-Guys-T2 veri setlerinde sırasıyla 38,72 dB ve 37,04 dB PSNR değerlerine ve 0,82 ve 0,81 SSIM değerlerine ulaşmış olup, bu da modelin görüntü kalitesini korumadaki güçlü kabiliyetini göstermektedir. Bu çalışma, DL modellerine kalıntı bağlantılar eklemenin, MRI görüntülerinden gürültüyü giderme yeteneklerini ve klinik ortamlarda tıbbi görüntülerin bütünlüğünü korumak için bir çözüm sunduğu sonucuna varmıştır.

Anahtar Kelimeler: Gürültü giderimi, Manyetik rezonans görüntüleme, Derin öğrenme, Artık bağlantı, UNet

Cite

Karakis, R., Topdag, T., (2024). "Noise Removal in Magnetic Resonance Imaging Using 3D Deep Learning Model", *Mugla Journal of Science and Technology*, 10(2), 31-41.

1. Introduction

Medical imaging techniques are crucial for the early diagnosis of diseases and the effective execution of treatments. Various techniques are used to visualize structures and functions within the brain, with Magnetic Resonance Imaging (MRI) being one of the most commonly employed methods [1-4]. Noise removal is a fundamental step in numerous computer vision applications, including image processing and medical imaging. In MRI, various types and levels of noise can occur during image acquisition or transmission, which may decrease image quality and cause distortion. For this reason, noise can impede the accurate interpretation of the disease and the determination of appropriate treatment.

In the literature, many methods have been proposed to remove noise in digital images [5]. Noise removal techniques can typically be categorized into conventional methods and those based on deep learning (DL), which is the focus of this study. Local smoothing methods in the spatial plane and basic filters in the frequency plane are commonly used for noise removal. However, these traditional filters cannot preserve the fine structures, details, and textures in the image, which can lead to information loss in MRI images. For this reason, Buades and colleagues [6] developed the Non-Local Means (NLM) filter, which takes advantage of the redundancy of information in the image. Dabov et al. [7] introduced the block-matching and 3D filtering (BM3D) method, which is an adaptation of the NLM algorithm. The performance of the BM3D filter decreases when the noise in the image gradually increases, and distortions may occur after filtering, especially in flat regions. Manjón et al [8] adapted the NLM filter to remove noise in MRI images by tuning the optimum parameters of the filter. In their study, Krissian and Aja-Fernández [9] developed a method based on estimating the standard deviation of noise to remove noise. This method combines local linear minimum mean square error filters applied to the Rician noise distribution with the anisotropic diffusion filter proposed by Perona and Malik [10]. While the noise removal of the method is successful it requires tuning of parameters such as the standard deviation of the Gaussian kernel used in filtering.

Traditional methods used to remove noise in MRI images have some challenges and parameters that need to be tuned. These challenges include determining the optimal window size, avoiding unwanted residuals after noise removal, low success in high-noise images, lack of detail preservation, and difficulty in parameter optimization [11]. To overcome these challenges, DL-based noise removal methods have been proposed in recent years. DL techniques consume fewer resources, can adapt to variable noise types, and are easy to use in clinical settings [12].

When examining the DL models proposed for noise removal in the literature [13], they can be categorized

into Convolutional Neural Network (CNN), Generative Adversarial Network (GAN), and autoencoder-based models.

CNN-based denoising methods are used to remove different types of noise, especially Gaussian and Rician, using CNN, which is one of the widely used DL models in image analysis. Zhang et al. [14] proposed the Denoising CNN (DnCNN) model to reduce Gaussian noise. In the model, batch normalization (BN) was used at the output of each layer, and performance was improved through residual learning and regularization. Jiang et al. [15] proposed two DL models consisting of multi-channel and residual-connected convolutional blocks to reduce Rician noise in MRI images. By adapting the DnCNN model to MRI images, MCDnCNNg models were developed for images with unknown noise levels, and MCDnCNNs models were developed for images with known noise levels. Li et al. [16] reduced noise by combining two ResNet models in the RicianNet model, where the first ResNet model lacked BN blocks while the second had them. Wu et al. [17] introduced the 3DParallel-RicianNet model, incorporating residual dilated convolution (DCR) and residual depthwise separable convolution (DSCR) modules to prevent volumetric feature loss and extract more local information from brain structures. In this context, the reviewed studies demonstrate the significant benefits of residual connections and BN layers in noise removal.

GAN-based denoising methods utilize the GAN model, which relies on the competition between a generator and a discriminator to produce clean images from noisy input images. Li et al. [18] suggested the conditional GAN (CGAN) model. Input images of the generator were obtained by adding gradient information to the noise vector. Ran et al. [19] proposed the 3D residual-connected Wasserstein GAN (RED-WGAN) architecture for reducing Rician noise in MRI images while preserving structures in the images. The model employed an optimization function that combines mean squared error (MSE), VGG, and Wasserstein errors for updating weight and bias values. When examining GAN-based denoising methods, challenges such as mode collapse and vanishing gradients between the generator and discriminator networks can occur during the learning process which may result in noisy images.

Autoencoder-based denoising methods rely on autoencoder models, which compress the input image to obtain high-level features and then aim to reconstruct clean images using these features. The autoencoder model, which is the focus of this study, is widely used in noise removal while preserving structural features and shows successful performance. Gondara [20] reduced noise in medical images using the CNN-DAE model containing convolutional blocks and achieved high SSIM values. Bermudez et al. [21] removed Gaussian noise in 2D MRI images with a residual-linked autoencoder network and achieved better performance than the SUSAN technique in FSL (FMRIB Software Library).

Tripathi et al. [22] proposed the CNN-DMRI model containing residual links between the encoder and decoder to reduce Rician noise and obtained high PSNR values for a 15% noise rate. Yang et al. [23] proposed a DL model based on Multilayer Perceptron (MLP) and autoencoder to reduce noise in 3D MRI images with small lesions and obtained high-performance results.

In this study, we aimed to eliminate Rician noise from MRI images, facilitating precise medical image analysis in clinical contexts and ensuring accurate disease diagnosis and treatment. For this reason, a dataset was first created by adding noise at different levels to T1 and T2 weighted MRI images from the IXI dataset (<http://brain-development.org/ixi-dataset/>). This dataset was used to train and test the two modified DL models developed by adding residual and dense connections to the UNet model in the study. The performance of DL models was evaluated with the structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), and mean absolute error (MAE) metrics.

The remainder of the paper is structured as follows: Section 2 explains the materials and the proposed deep-learning methods. Section 3 presents the results obtained from the proposed methods and discusses these findings compared with similar studies in the literature. Finally, Section 4 covers the conclusions and potential directions for future research.

2. Material and Methods

In this section, the details of two DL models that are modified from the 3D UNet architecture for noise removal in 3D MRI images are provided. Additionally, information regarding the dataset used for training and testing these models is presented. This section also includes a comprehensive explanation of the performance metrics utilized to evaluate the models, as well as details regarding the Rician noise applied to the MRI images.

2.1. Dataset

The publicly available IXI dataset (<http://brain-development.org/ixi-dataset/>) was used for training and testing the developed DL models. The IXI dataset contains approximately 600 MRI images of healthy individuals in T1 and T2 weighted NIFTI(.nii) format, gathered from three different hospitals (Guys, HH: Hammersmith, and IOP: Institute of Psychiatry). The number of IOP samples in the dataset is limited and therefore cannot meet the data requirement for generalizing a problem in the DL model. For this reason, following similar approaches in the literature, DL models were trained and tested separately only the T1 and T2 weighted 3D MRI images from Guys and HH hospitals.

The T1 and T2-weighted MRI images were divided into 80% for training and 20% for testing. The original images served as the ground truth for the output of the DL model. The resolutions of images in the dataset vary, with dimensions of 256×256×150 and 256×256×146. To

ensure compatibility with DL models, all images underwent preprocessing during runtime, involving zero-padding to standardize the resolution to 256×256×160. The noisy images, which are the input to the network, were created by adding Rician noise at rates of 3%, 5%, 7%, 9%, 11%, 13%, and 15% to the original images. Rician noise addition was implemented using the `add_noise()` function available in the `dipy` library (https://dipy.org/documentation/1.1.1./reference/dipy.sims/#dipy.sims.voxel.add_noise).

2.2. Rician Noise

MRI images contain varying levels of noise caused by a range of factors, including random fluctuations, physiological processes, differences in magnetic susceptibilities between adjacent tissues, body movements, and environmental conditions. Different types of noise, such as Rician noise in MRI, can distort image contrast during image acquisition or transmission, leading to erroneous diagnoses [24]. Rician noise refers to the discrepancy between the actual image intensities and the observed data. [25-26]. Rician noise in the MR image is as shown in Equation (1).

$$Y = A + N \quad (1)$$

Where Y is the image with added noise, A is the original image, and N denotes the Rician noise added to the original image. The primary purpose of denoising is to remove Rician noise, particularly preserving structural details with minimal impact on the original image (A). The Rician noise is modeled as shown in Equation (2) [26].

$$p_M(M) = \frac{M}{\sigma^2} e^{-(M^2-A^2)/2\sigma^2} I_0\left(\frac{AM}{\sigma^2}\right) \quad (2)$$

In Equation 2, I_0 denotes the zero-order Bessel function, σ is the standard deviation of Gaussian noise, M represents the relevant pixel value of the noisy image, and A is the pixel value in the original image at the same location.

2.3. 3D Deep Learning Model

The UNet architecture consists of the encoder segment, which condenses the input image into smaller feature maps, and the decoder segment, which converts these feature maps back into the input image dimensions, as shown in Fig. 1. Residual connections bridging the two segments facilitate the network in capturing finer details. Fig. 1 illustrates the general framework of the UNet architecture [27].

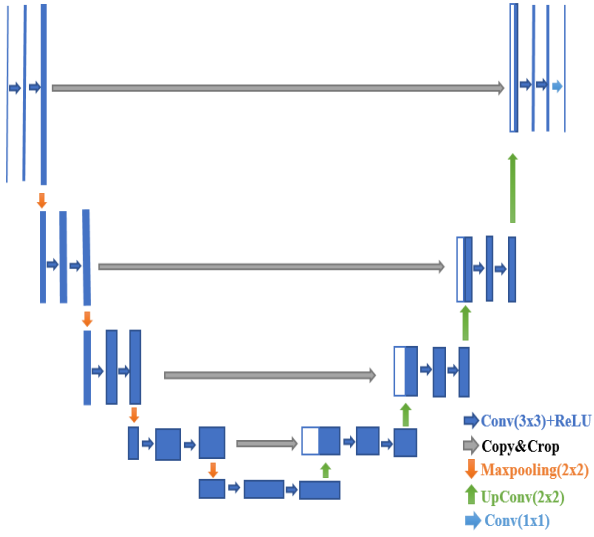


Figure 1. The framework of the UNet architecture [27]

In this study, a DL model was proposed for removing Rician noise in MRI images using the UNet architecture, known for denoising, segmentation, and image enhancement [28]. The UNet model was modified to operate on 3D volumetric MRI images. Fig. 2 illustrates the 3D architecture of the proposed model, named M-UNet. In the M-UNet architecture, ResNet50 [29] and DenseNet121 [30] layers were inserted into the latent layer between the encoder and decoder blocks of the traditional UNet, allowing for more compression of features with residual or dense connections [31-35].

As shown in Fig. 2, the encoder section of the M-UNet model comprises four 3D encoder layers. Each of these blocks executes 3D convolution, batch normalization (BN), Leaky ReLU activation, and max pooling operations. The convolution operation is employed to detect features in the image using filters. This process involves sliding filters of various sizes (e.g., 3x3, 5x5, 7x7) across all pixels of the image, resulting in the creation of feature maps. Feature maps are regions where specific characteristics of the filter are identified. BN is utilized to normalize the values within the feature maps, thereby enhancing their consistency. Non-linear activation functions (such as Leaky ReLU, ReLU, sigmoid, softmax, etc.) are essential in deep learning to introduce non-linear real-world features to the network. ReLU zeros out the negative values of neurons while retaining the positive values as they are. In contrast, Leaky ReLU preserves the negative values by multiplying them with a small slope while retaining the positive values as they are. Leaky ReLU addresses the dead neuron and gradient vanishing problems associated with negative values in ReLU. The pooling layer reduces the dimensions of the feature maps, thus decreasing the number of parameters. The primary goal here is to preserve significant information while reducing the size. In commonly used max pooling and average pooling, the maximum or average value within the window slid over the feature map is taken, respectively, to reduce the size [28].

As shown in Fig. 2, the decoder section of the M-UNet model consists of four 3D decoding layers. Each of these layers performs 3D transposed convolution (Conv3DTranspose), the addition of features from the encoding block, 3D convolution, batch normalization (BN), and Leaky ReLU operations. The transposed convolution layer enables the reconstruction of higher-resolution features from the encoded features by multiplying the input data with filter values. The addition layer combines the features from the encoding layer, received via skip connections, with those obtained through transposed convolution, thereby optimizing the gradient flow between layers. In the final decoder layer, a single-filter 3D convolution is applied with a sigmoid activation function, resulting in an output image that matches the volumetric dimensions of the input image [28].

2.4. Performance Metrics

In the study, the similarity between the real images and the noise-free images obtained by the DL model was determined using PSNR, SSIM, and MAE metrics.

PSNR is determined based on the relationship between the maximum value of the image and MSE, as outlined in Equation (1).

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (1)$$

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

Where I and K represent the original and the noise-free image, respectively, while m and n denote the width and height of the images.

SSIM determines the structural similarity and perceptual image quality between images by comparing their luminance, contrast, and structural features, as described in Equation (2).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

Here x and y are the images being compared, and the values μ_x , μ_y , σ_x , σ_y , and σ_{xy} represent the mean of pixel intensities, standard deviation, and covariance of the images, respectively.

MAE is obtained by calculating the absolute difference between the predicted and the actual values, as detailed in Equation (3).

$$MAE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)| \quad (3)$$

3. Experimental Results

In the IXI open dataset, 3D T1 and T2 MRI images from Guys and HH hospitals were divided into training and testing datasets with an 80% to 20% split for the DL models. The testing data were also utilized as validation data to prevent the networks from overfitting. Rician noise was added to the images at levels of 3%, 5%, 7%, 9%, 11%, 13%, and 15%. The noisy images served as inputs, while the original images served as outputs for the DL models.

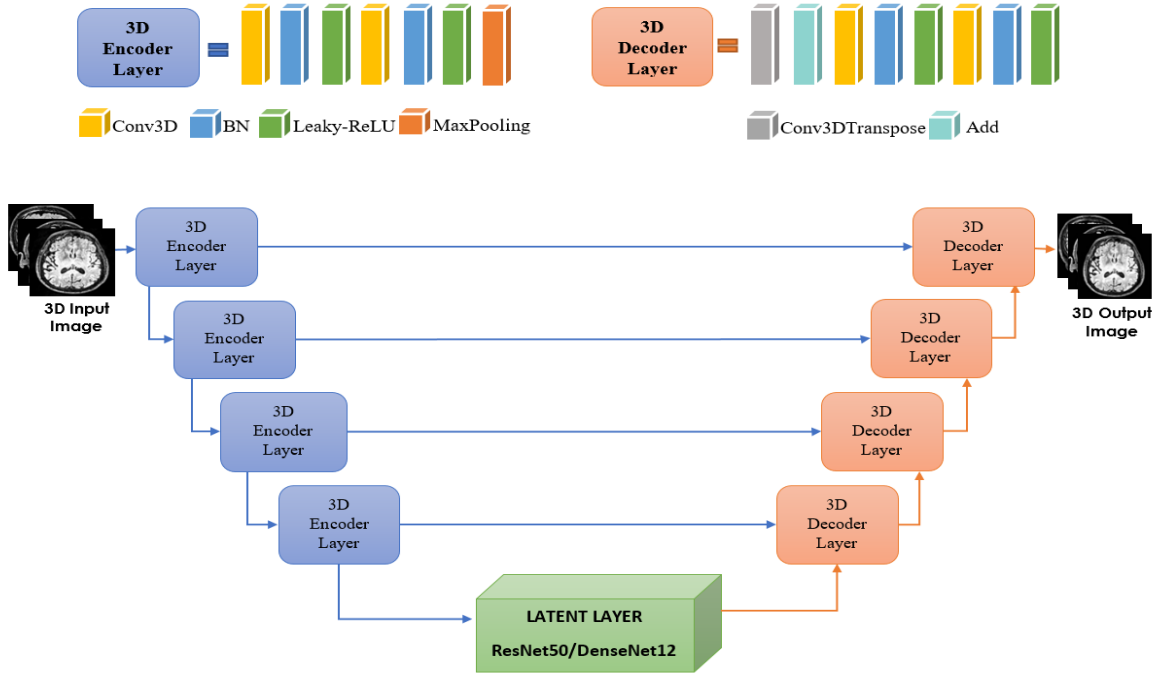


Figure 2. Proposed 3D M-UNet architecture for noise-removal [28].

The models were implemented and analyzed using Python 3.9, along with the Keras and Tensorflow libraries. The analysis of the BM4D filter [7, 36] was conducted using the MATLAB development environment. The experiments were carried out on a system equipped with an NVIDIA RTX A6000 48 GB GPU, an Intel i9-12900KS @ 3.40 GHz CPU, and 64 GB of RAM. The Adam optimization algorithm, in conjunction with dropout and L2 regularization, was used to update the free parameters, including weights and biases, and to prevent overfitting. The learning rate, momentum coefficient, and weight decay were set to 0.001, 0.8, and 0.00001, respectively. The models were trained for 50 epochs with a batch size of 4.

The following subsections present the performance of 3D UNet-based DL models in denoising T1 and T2-weighted images from the IXI-Guys and IXI-HH datasets, across varying levels of noise. Additionally, a discussion is provided on the comparison of these models' performance with similar studies in the literature, highlighting both their strengths and limitations in the context of noise reduction.

3.1. Results for the IXI-Guys Dataset

Table 1 presents the performance results of the M-UNet, which has been modified with ResNet and DenseNet in its latent layers, for noisy T1 and T2-weighted test MRI images from the IXI-Guys dataset. For T1-weighted images, the M-UNet+ResNet model's PSNR values ranged from 40.52 at 3% noise to 34.17 at 15% noise, while SSIM values decreased from 0.95 to 0.88. Similarly, the M-

UNet+DenseNet model showed a decline in PSNR from 40.23 to 34.26 and SSIM from 0.95 to 0.85.

In T2-weighted images, the M-UNet+ResNet model demonstrated a higher starting PSNR of 47.90 at 3% noise, decreasing to 38.72 at 15% noise, with SSIM values dropping from 0.95 to 0.82. The M-UNet+DenseNet model followed a similar trend, with PSNR values decreasing from 47.37 to 38.20 and SSIM from 0.95 to 0.85. Overall, T2-weighted images exhibited higher initial PSNR and SSIM values compared to T1-weighted images, indicating potentially better baseline image quality. Additionally, the M-UNet+ResNet model generally outperformed the M-UNet+DenseNet model in maintaining higher PSNR and SSIM values across varying noise levels, though the differences were modest.

Fig. 3 shows the loss values (binary cross-entropy) calculated by the DL models for different noise levels. As observed in Figures 3, the loss values of the DL models increased as the noise level rises. Fig. 4 illustrates the axial slice of a T2-weighted test MRI image from the IXI-Guys dataset along with images with added noise levels of 5%, 9%, and 15%. It also includes the predicted images obtained using the M-UNet+ResNet and M-UNet+DenseNet models. The results indicated that while the performance of both models in removing different noise levels was comparable, the M-UNet+ResNet model performed slightly better overall.

Table 1. Performance results for noisy T1 & T2-weighted test MRI images from the IXI-Guys dataset.

Noise Ratio	T1-weighted MRI images						T2 weighted MRI images					
	M-UNet+ResNet			M-UNet+DenseNet			M-UNet+ResNet			M-UNet+DenseNet		
	PSNR(dB*)	SSIM	MAE	PSNR	SSIM	MAE	PSNR	SSIM	MAE	PSNR	SSIM	MAE
3%	40.52	0.95	0.007	40.23	0.95	0.008	47.90	0.95	0.005	47.37	0.95	0.005
5%	37.66	0.90	0.010	37.74	0.94	0.010	42.92	0.91	0.008	43.64	0.91	0.008
7%	36.78	0.90	0.011	37.16	0.92	0.011	42.78	0.92	0.009	40.33	0.88	0.011
9%	36.47	0.92	0.012	35.16	0.90	0.014	41.47	0.88	0.011	40.83	0.87	0.012
11%	35.64	0.91	0.013	35.21	0.90	0.013	41.07	0.88	0.013	39.77	0.83	0.015
13%	34.57	0.88	0.015	34.69	0.89	0.014	40.46	0.87	0.015	39.64	0.84	0.016
15%	34.17	0.88	0.015	34.26	0.88	0.015	38.72	0.82	0.018	38.20	0.85	0.017
Average	36.54	0.91	0.012	36.35	0.91	0.012	42.19	0.89	0.011	41.40	0.88	0.012
SD**	2.14	0.02	0.002	2.13	0.02	0.002	2.89	0.04	0.004	3.11	0.04	0.004

*dB: decibel, **SD: standard deviation

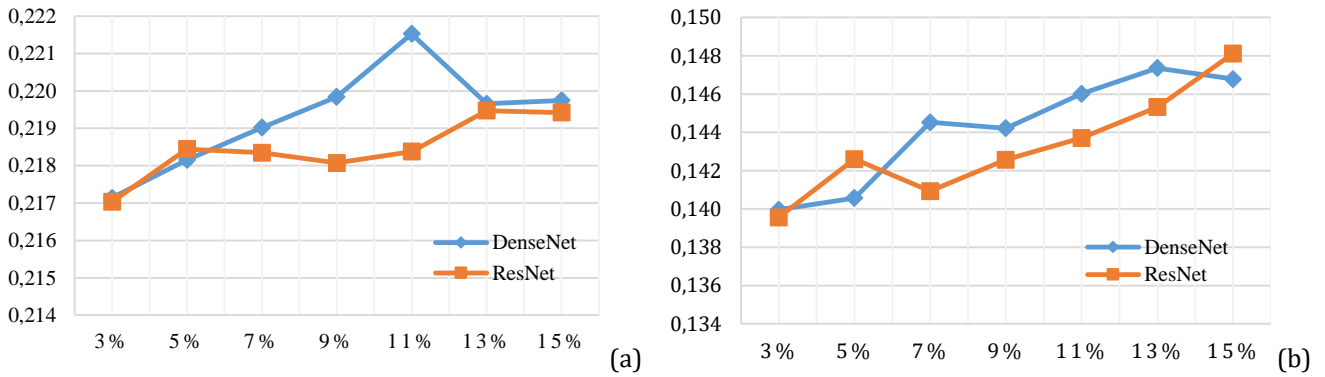


Figure 3. Loss values computed for T1 & T2-weighted MRI images in the IXI-Guys dataset.

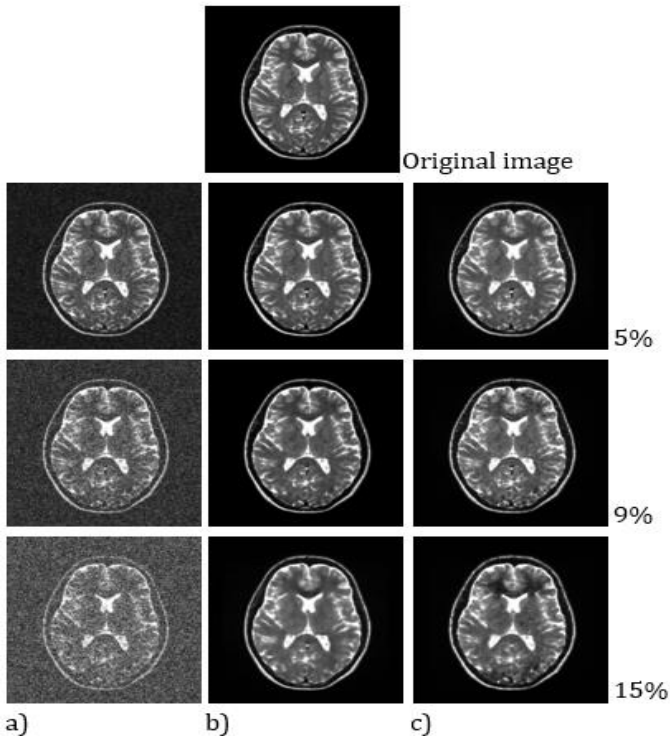


Figure 4. Original T2-weighted MRI axial slice from the IXI-Guys dataset, (a) noisy image, (b) M-UNet+ResNet prediction, (c) M-UNet+DenseNet prediction.

3.2. Results for the IXI-HH Dataset

Table 2 provides the performance metrics for the M-UNet models on noisy T1 and T2-weighted test MRI images from the IXI-HH dataset. As the noise ratio increases, both models exhibited a decline in PSNR and SSIM values, indicating a reduction in image quality, while MAE values increased, reflecting higher prediction errors.

For T1-weighted images, the PSNR for the M-UNet+ResNet model decreased from 42.35 dB to 34.74 dB and the SSIM from 0.96 to 0.85, whereas the M-UNet+DenseNet model showed similar trends with PSNR dropping from 42.48 dB to 34.85 dB and SSIM from 0.96 to 0.84. In the case of T2-weighted images, the M-UNet+ResNet model's PSNR decreased from 46.71 dB to 37.04 dB and SSIM from 0.96 to 0.81, while the M-UNet+DenseNet model's PSNR dropped from 46.72 dB to 34.93 dB and SSIM from 0.96 to 0.78.

Figure 5 illustrates the loss values calculated by the models across various noise levels, showing that these values generally escalate as the noise level intensifies. It was observed that the error values produced by the models were particularly high for the noise levels of 13% and 15%, compared to the other noise levels.

Table 2. Performance results for noisy T1 & T2-weighted test MRI images from the IXI-HH dataset.

Noise Ratio	T1-weighted MRI images						T2-weighted MRI images					
	M-UNet+ResNet			M-UNet+DenseNet			M-UNet+ResNet			M-UNet+DenseNet		
	PSNR(dB*)	SSIM	MAE	PSNR	SSIM	MAE	PSNR	SSIM	MAE	PSNR	SSIM	MAE
3%	42.35	0.96	0.005	42.48	0.96	0.005	46.71	0.96	0.004	46.72	0.96	0.004
5%	39.78	0.93	0.008	39.90	0.93	0.008	45.80	0.94	0.006	44.68	0.94	0.006
7%	38.76	0.92	0.009	38.53	0.92	0.009	42.23	0.91	0.008	41.69	0.88	0.009
9%	37.23	0.87	0.011	37.56	0.89	0.010	40.86	0.90	0.009	40.68	0.88	0.011
11%	35.82	0.88	0.012	37.27	0.90	0.010	40.34	0.88	0.011	38.34	0.85	0.013
13%	36.18	0.87	0.012	35.76	0.87	0.012	39.19	0.88	0.012	37.35	0.82	0.017
15%	34.74	0.85	0.014	34.85	0.84	0.015	37.04	0.81	0.017	34.93	0.78	0.020
Mean	37.84	0.90	0.010	38.05	0.90	0.010	41.74	0.90	0.009	40.63	0.87	0.011
SD**	2.63	0.03	0.003	2.56	0.00	0.003	3.48	0.04	0.004	4.14	0.06	0.005

*dB: decibel, **SD: standard deviation

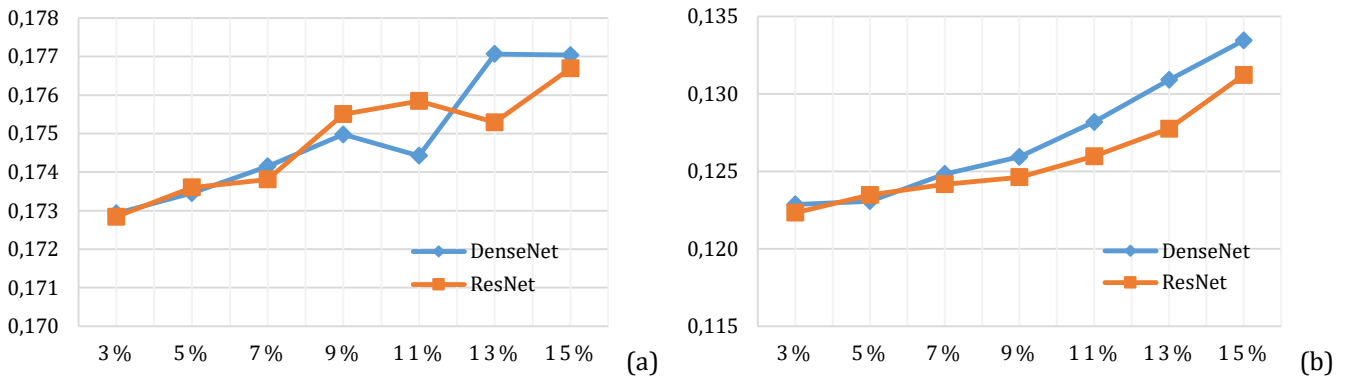


Figure 5. Loss values computed for T1 & T2 weighted MRI images in the IXI-HH dataset.

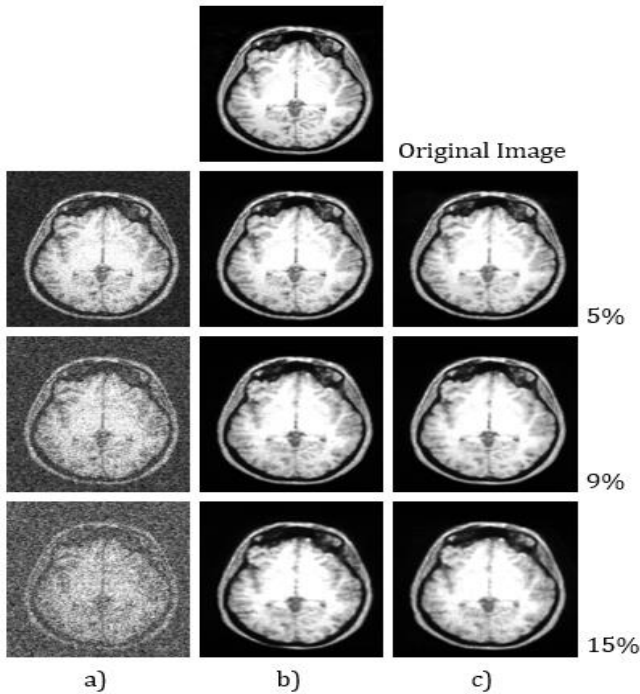
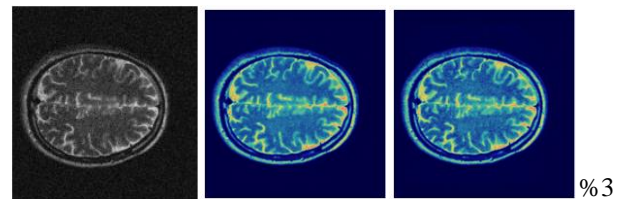


Figure 6. Original T1-weighted MRI axial slice from the IXI-HH, (a) noisy images, (b) M-UNet+ResNet prediction, (c) M-UNet+DenseNet prediction.

Figure 6 shows the axial slice of a T1-weighted test MRI image from the IXI-HH dataset, along with images with

added noise levels of 5%, 9%, and 15%, and the corresponding predictions obtained using M-UNet+ResNet and M-UNet+DenseNet. Overall, the M-UNet+ResNet model had slightly better image quality metrics than the M-UNet+DenseNet model, despite both showing a similar decline in performance as noise levels increase.

In Figure 7, Grad-CAM activation maps were provided to demonstrate the models' explainability between the predicted and ground truth outputs. As noise levels increased, the results from DenseNet showed less effective noise targeting, with more diffused activations. In contrast, at lower noise levels, both models produced comparable results. M-UNet+ResNet generally showed more concentrated and focused activations, targeting noise effectively in specific regions while preserving structural details. This indicated its superior performance in denoising and maintaining image quality. On the other hand, M-UNet+DenseNet had a broader approach with less localized noise reduction.



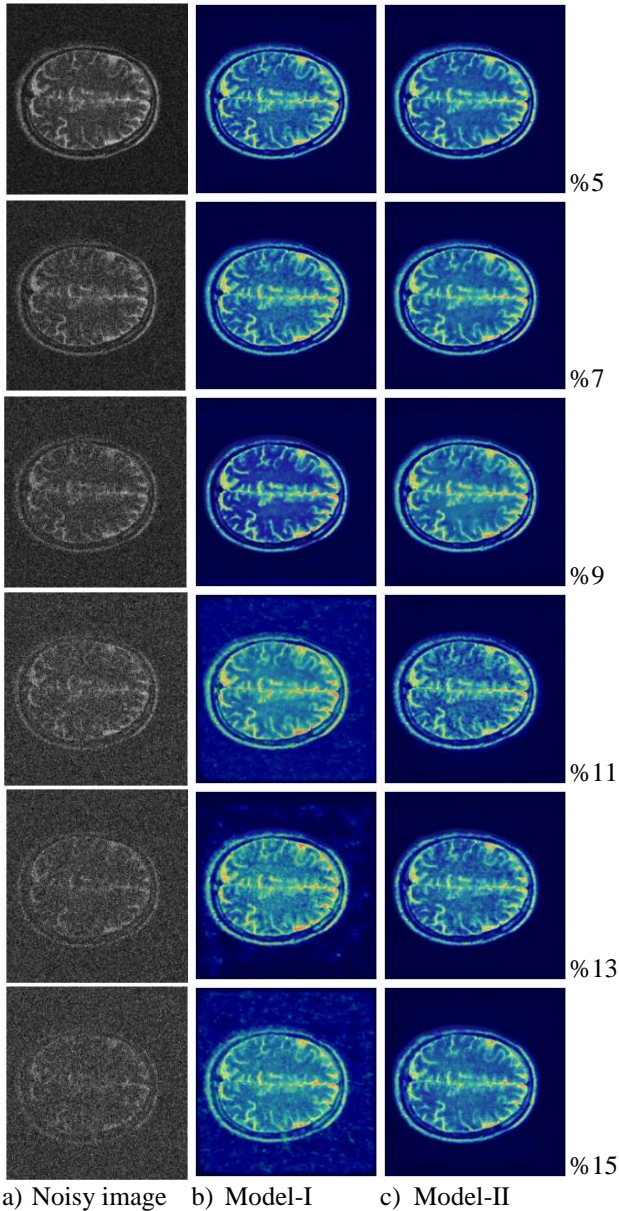


Figure 7. Grad-CAM activation maps of DL models for various noise levels, a) noisy image, b) M-UNet+DenseNet, c) M-UNet+ResNet.

3.3. Literature Comparison

Table 3 compares the PSNR values obtained for noise removal using different DL models on the IXI dataset, including the proposed M-UNet+ResNet and M-UNet+DenseNet models. As the noise levels in the images increase, particularly at 11%, 13%, and 15%, the performance metrics decrease compared to lower noise levels. This is because higher noise levels make it difficult to distinguish the structures in the images from the noise, leading to significant losses in essential structures during the noise removal process. As seen in Table 3, the highest PSNR results reported in the literature were achieved using the MCDnCNNs model proposed by Zhang et al. [14], with a value of 31.62 dB for the 15% noise level. The RED-WGAN model [19], tested on both the IXI-HH-T1

and IXI-HH-T2 datasets, achieves PSNR values of only 29.55 dB and 29.79 dB at 15% noise.

In contrast, the proposed M-UNet+ResNet model showed PSNR values between 42.35 dB and 34.74 dB on the HH-T1 dataset, and between 46.71 dB and 37.04 dB on the HH-T2 dataset. This indicated strong performance, particularly for lower noise levels, with a slight decline as noise increased. Similarly, the M-UNet+DenseNet model performed comparably, with PSNR values ranging from 42.48 dB to 34.85 dB for HH-T1 and from 46.72 dB to 34.94 dB for HH-T2. In addition, the proposed models have achieved better results compared to the traditional 3D-UNet in terms of PSNR values at various noise levels, as can be seen in Table 3.

Table 3. Comparison of PSNR values for the proposed DL models with existing literature

DL model	DS*	3%	5%	7%	9%	11%	13%	15%
MCDnCNNg [14]	HH-T1	39.38	37.12	35.40	33.86	32.54	31.10	29.96
MCDnCNNs [14]	HH-T1	40.47	37.82	36.20	34.71	33.56	32.57	31.62
RED-WGAN [19]	HH-T1	36.52	34.46	33.03	33.03	32.14	30.60	29.55
RED-WGAN [19]	HH-T2	38.57	36.22	34.57	33.09	31.91	30.76	29.79
UNET-3D	HH-T1	42.02	39.90	38.34	37.69	37.20	34.14	34.08
UNET-3D	HH-T2	46.42	45.31	42.67	39.61	40.22	38.52	34.82
M-UNET+ResNet	HH-T1	42.35	39.78	38.76	37.23	35.83	36.18	34.74
M-UNET+DenseNet	HH-T1	42.48	39.90	38.53	37.56	37.27	35.77	34.85
M-UNET+ResNet	HH-T2	46.71	45.80	42.23	40.86	40.34	39.19	37.04
M-UNET+DenseNet	HH-T2	46.72	44.68	41.69	40.68	38.34	37.35	34.94

*DS: dataset

Table 4 presents the performance results of the M-UNET models compared to the BM4D filter [7, 36], which is commonly used for noise removal in MRI images. The analyses were conducted under the same conditions using the test data employed for evaluating the DL models. According to Table 4, both proposed models outperformed the BM4D filter in terms of PSNR values. According to Tables 3 and 4, the M-UNET+ResNet and M-UNET+DenseNet models, which have been modified with residual connections, demonstrate generally high performance in noise removal. In this context, the results indicated that the M-UNET+ResNet model, in particular, outperformed existing methods in the literature, effectively preserving structural details in the images.

Table 4. Comparison of PSNR values between the BM4D filter and the proposed DL models

DS*	Method	3%	5%	7%	9%	11%	13%	15%	Mean	SD**
Guys-T1	BM4D	31.43	27.71	24.99	22.88	21.15	19.68	18.41	23.75	4.63
	M-UNET+ResNet	40.52	37.66	36.78	36.47	35.64	34.57	34.17	36.54	2.14
	M-UNET+DenseNet	40.23	37.74	37.16	35.16	35.21	34.69	34.26	36.35	2.13
Guys-T2	BM4D	31.57	27.58	24.74	22.55	20.77	19.28	17.99	23.50	22.34
	M-UNET+ResNet	47.90	42.92	42.78	41.47	41.07	40.46	38.72	42.19	2.89
	M-UNET+DenseNet	47.37	43.64	40.33	40.83	39.77	39.64	38.20	41.40	3.11
HH-T1	BM4D	32.03	27.86	24.97	22.76	20.97	19.46	18.17	23.75	4.92
	M-UNET+ResNet	42.35	39.78	38.76	37.23	35.83	36.18	34.74	37.84	2.63
	M-UNET+DenseNet	42.48	39.90	38.53	37.56	37.27	35.76	34.85	38.05	2.56
HH-T2	BM4D	32.09	27.73	24.76	22.51	20.70	19.18	17.87	23.55	5.04
	M-UNET+ResNet	46.71	45.80	42.23	40.86	40.34	39.19	37.04	41.74	3.48
	M-UNET+DenseNet	46.72	44.68	41.69	40.68	38.34	37.35	34.93	40.63	4.14

*DS: dataset, **SD= standard deviation

The comparison of models in terms of average prediction time was performed on the HH-T2 test data. Accordingly, the number of trainable parameters for the M-UNET+ResNet and M-UNET+DenseNet models are 2,088,925 and 6,920,009, respectively. The average prediction times for the M-UNET+ResNet, M-UNET+DenseNet models, and the BM4D filter are 3.10 seconds (± 0.04), 4.02 seconds (± 0.05), and 21.80 seconds (± 0.67), respectively. The M-UNET+ResNet model is more efficient than the M-UNET+DenseNet and the BM4D filter, offering faster prediction times and fewer trainable parameters, making it the better choice for computationally efficient applications.

Although the developed model demonstrated high performance, the study has certain limitations. First, the experiments used the IXI dataset, a robust and diverse dataset containing various T1 and T2-weighted MRI images from different hospitals and devices to enhance the model's generalization capability. However, this choice may still limit the generalizability of the results when applying the model to datasets with different acquisition parameters. Second, the models were specifically designed for 3D MRI images, leaving their performance on 2D MRI images or other imaging modalities unexplored. However, it should be noted that 3D models can achieve better performance due to the benefit of neighborhood relationships in the depth direction. Additionally, the denoising performance may vary when applied to MRI images obtained under different scanning conditions or noise characteristics, as the models were trained using synthetic noise levels. While BM4D was used as a baseline filter for comparison,

it is important to acknowledge that every filtering method has its limitations. However, as shown in Table 4, the proposed models demonstrated notable effectiveness in preserving structural details and reducing noise, highlighting their potential to address some of the challenges associated with traditional noise removal techniques.

4. Conclusion

Noise reduction from medical images has become recently essential for the accurate diagnosis of diseases. This study discusses traditional and DL methods for noise removal and evaluates the success of developed DL models. A DL architecture similar to the UNet was designed, featuring an autoencoder structure with the encoder, decoder, and middle latent layers containing dense and residual layers from pre-trained ResNet50 and DenseNet121 models. These modifications resulted in the creation of two methods, M-UNET+ResNet and M-UNET+DenseNet. The performance of these proposed models was evaluated across different noise levels using quality metrics such as PSNR, SSIM, and MAE, and they were compared to methods found in the literature. The proposed models were trained to remove varying noise levels, ranging from 3% to 15%, added to the IXI-Guys and IXI-HH T1 and T2 MRI images. Generally, M-UNET+ResNet showed superior results compared to M-UNET+DenseNet in IXI-HH-T2 and IXI-Guys-T2 datasets, while both models produced similar results in IXI-HH-T1 and IXI-Guys-T1 datasets. T2-weighted images performed better than T1-weighted images overall. This may be due to the different characteristics and applications of T1 and T2-weighted images, where T1-weighted images provide high soft tissue contrast and spatial resolution, useful for anatomical details, while T2-weighted images highlight pathological changes and fluids in the brain. The proposed architectures achieved higher PSNR and SSIM values than existing noise removal approaches in the literature, indicating that DL models with residual connections can significantly enhance performance in noise removal. Future research could explore the success of specially designed DL models with residual connections and custom blocks for noise removal, using different loss functions to better preserve structural details. Additionally, since patient data were not used in this study, the effectiveness of noise removal could be tested on medical images from patients with lesions.

5. Acknowledgment

This work is the result of the Master of Science (MS) thesis study of the second author.

6. References

- [1] Oyar, O., "Magnetik Rezonans Görüntüleme MRG'nin klinik uygulamaları ve endikasyonları", *Harran Üniversitesi Tıp Fakültesi Dergisi*, Vol. 5, No. 2, 31-40, 2008.

- [2] Gürkahraman, K., and Karakiş, R., "Brain tumors classification with deep learning using data augmentation", *Journal of the Faculty of Engineering and Architecture of Gazi University*, Vol. 36, No. 2, 997-1011, 2021.
- [3] Yapıcı, M., Karakiş, R., and Gürkahraman, K., "Improving brain tumor classification with deep learning using synthetic data", *Computers, Materials and Continua*, Vol. 74, No. 3, 2023.
- [4] Karakis, R., Gurkahraman, K., Mitsis, G. D., Boudrias, M. H., "Deep learning prediction of motor performance in stroke individuals using neuroimaging data", *Journal of Biomedical Informatics*, Vol. 141, article id: 104357, 2023.
- [5] Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., Lin, C. W., "Deep learning on image denoising: An overview", *Neural Networks*, Vol. 131, 251-275, 2020.
- [6] Buades, A., Coll, B., Morel, J. M., "A review of image denoising algorithms, with a new one". *Multiscale modeling & simulation*, Vol. 4, No. 2, 490-530, 2005.
- [7] Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K. "Image denoising by sparse 3-D transform-domain collaborative filtering", *IEEE Transactions on image processing*, Vol. 16, No. 8, 2080-2095, 2007.
- [8] Manjón, J. V., Carbonell-Caballero, J., Lull, J. J., García-Martí, G., Martí-Bonmatí, L., Robles, M., "MRI denoising using non-local means", *Medical image analysis*, Vol. 12, No. 4, 514-523, 2008.
- [9] Krissian, K., and Aja-Fernández, S., "Noise-driven anisotropic diffusion filtering of MRI", *IEEE transactions on image processing*, Vol. 18, No. 10, 2265-2274, 2009.
- [10] Perona, P., and Malik, J., "Scale-space and edge detection using anisotropic diffusion", *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 12, No. 7, 629-639, 1990.
- [11] Patil, R., and Bhosale, S., "Medical Image Denoising Techniques: A Review", *International Journal on Engineering, Science and Technology (IJonEST)*, Vol. 4, No. 1, 21-33, 2022.
- [12] Lundervold, A. S., and Lundervold, A. "An overview of deep learning in medical imaging focusing on MRI", *Zeitschrift für Medizinische Physik*, Vol. 29, No. 2, 102-127, 2019.
- [13] Ilesanmi, A. E., and Ilesanmi, T. O., "Methods for image denoising using convolutional neural network: a review", *Complex & Intelligent Systems*, Vol. 7, No. 5, 2179-2198, 2021.
- [14] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., "Beyond a gaussian denoiser: Residual learning of deep cnn for image Denoising", *IEEE Transactions on Image Processing*, Vol. 26, No. 7, 3142-3155, 2017.
- [15] Jiang, D., Dou, W., Vosters, L., Xu, X., Sun, Y., Tan, T., "Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network", *Japanese journal of radiology*, Vol. 36, No. 9, 566-574, 2018.
- [16] Li, S., Zhou, J., Liang, D., & Liu, Q., "MRI denoising using progressively distribution-based neural network", *Magnetic resonance imaging*, Vol. 71, 55-68, 2020.
- [17] Wu, L., Hu, S., Liu, C., "Denoising of 3D Brain MR images with parallel residual learning of convolutional neural network using global and local feature extraction", *Computational Intelligence and Neuroscience*, Vol. 2021, article id: 5577956, 1-18, 2021.
- [18] Li, Y., Zhang, K., Shi, W., Miao, Y., Jiang, Z., "A Novel Medical Image Denoising Method Based on Conditional Generative Adversarial Network", *Computational and Mathematical Methods in Medicine*, Vol. 2021, article id: 9974017, 1-11, 2021.
- [19] Ran, M., Hu, J., Chen, Y., Chen, H., Sun, H., Zhou, J., Zhang, Y., "Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network", *Medical image analysis*, Vol. 55, 165-180, 2019.
- [20] Gondara, L., "Medical image denoising using convolutional denoising autoencoders", *2016 IEEE 16th international conference on data mining workshops (ICDMW)*, 241-246, 2016.
- [21] Bermudez, C., Plassard, A. J., Davis, L. T., Newton, A. T., Resnick, S. M., Landman, B. A., "Learning implicit brain MRI manifolds with deep learning", *In Medical Imaging 2018: Image Processing- SPIE*, 10574, 408-414, 2018.
- [22] Tripathi, P. C., and Bag, S., "CNN-DMRI: a convolutional neural network for denoising of magnetic resonance images", *Pattern Recognition Letters*, Vol. 135, 57-63, 2020.
- [23] Yang, H., Zhang, S., Han, X., Zhao, B., Ren, Y., Sheng, Y., Zhang, X. Y., "Denoising of 3D MR images using a voxel-wise hybrid residual MLP-CNN model to improve small lesion diagnostic confidence", *In International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, 292-302, 2022.
- [24] Savaji, S., and Arora, P., "Denoising of MRI images using thresholding techniques through wavelet", *International Journal of Innovative Science, Engineering & Technology*, Vol. 1, No. 7, 422-427, 2014.
- [25] Coupé, P., Manjón, J. V., Gedamu, E., Arnold, D., Robles, M., Collins, D. L. "Robust Rician noise estimation for MR images", *Medical image analysis*, Vol. 14, No. 4, 483-493, 2010.
- [26] Gudbjartsson, H., and Patz, S., "The Rician distribution of noisy MRI data", *Magnetic resonance in medicine*, Vol. 34, No. 6, 910-914, 1995.

- [27] Ronneberger, O., Fischer, P., Brox, T., “U-net: Convolutional networks for biomedical image segmentation”, In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, Springer International Publishing, 234-241, 2015.
- [28] Topdağ, T., “Beyin görüntülerindeki gürültülerin derin öğrenme kullanılarak giderilmesi”, Yüksek Lisans Tezi, *Sivas Cumhuriyet Üniversitesi Fen Bilimleri Enstitüsü*, Sivas, 2023.
- [29] He, K., Zhang, X., Ren, S., Sun, J., “Deep residual learning for image recognition”, In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778, 2016.
- [30] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q., “Densely connected convolutional networks”, In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700-4708, 2017.
- [31] Targ, S., Almeida, D., Lyman, K., “Resnet in resnet: Generalizing residual architectures”, *arXiv preprint, arXiv:1603.08029*, 2016.
- [32] Gurkahraman, K., and Daşgın, Ç., “Brain Extraction from Magnetic Resonance Images Using UNet modified with Residual and Dense Layers”, *Türk Doğa ve Fen Dergisi*, Vol. 12, No. 3, 144-151, 2023.
- [33] Savaş, S., “Detecting the stages of Alzheimer’s disease with pre-trained deep learning architectures”, *Arabian Journal for Science and Engineering*, Vol. 47, No. 2, 2201-2218, 2022.
- [34] Savaş, S., and Damar, Ç., “Transfer-learning-based classification of pathological brain magnetic resonance images”, *ETRI Journal*, Vol. 46, No. 2, 263-276, 2024.
- [35] Alhatemi, R. A. J., and Savaş, S., “A weighted ensemble approach with multiple pre-trained deep learning models for classification of stroke”, *Medinformatics*, Vol. 1, No. 1, 10-19, 2024.
- [36] Maggioni, M., Katkovnik, V., Egiazarian, K., Foi, A., “Nonlocal transform-domain filter for volumetric data denoising and reconstruction”, *IEEE transactions on image processing*, Vol. 22, No. 1, 119-133, 2012.