JEEEIT TRANSACTIONS, Vol.03,No.01,July 2026

Deep Learning-Based CT Image Denoising Using Internal Residual Learning with DnCNN

Farag J. Zbeda[1], Abdurahman E. Saleh[2] and Salem A. Salem[2]
[1]Computer Engineering Department, Azzaytuna University, Tarhuna, Libya.
[2]Computer Science Department, Higher Institute Awlad Ali, Tarhuna, Libya.

Corresponding author: (f_zbeda@yahoo.com).

ABSTRACT Image denoising is a fundamental task in image processing that aims to enhance the visual quality of images by removing unwanted noise introduced during acquisition, transmission, or storage. Noise can degrade visual content and affect image analysis tasks such as edge detection, object recognition, and feature extraction. The main challenge lies in balancing noise reduction with detail preservation. Traditional algorithms such as Gaussian and median filtering often blur important features of the images. This paper presents a deep learning scheme for denoising CT images, based on Internal Residual Learning with Deep Convolutional Neural Network (DnCNN). The model was used to suppress different levels of noise varying from 0.05 up to 0.4. The proposed model yielded very promising results compared to the classic Gaussian filter, Non-Local median (NLM), and standard DnCNN approaches.

KEYWORDS: image denoising, DnCNN, Internal Residual Learning, Gaussian filter, Non-Local median.

I. INTRODUCTION

The noise in images can arise from a variety of sources, such as sensor imperfections, low-light conditions and transmission errors, which degrade the visual quality of the images. This poses challenges for human perception as well as for automatic image analysis. Effective image denoising is very important for applications such as medical diagnosis, surveillance and remote sensing, where image clarity is very important. Traditional denoising methods such as Gaussian, median and non-local mean filtering can remove the noise, but they tend to blur the important structures and fine details of the image. [1] To overcome this problem, deep learning methods such as Deep Convolutional Neural Network (DnCNN) were used.

The DnCNN architecture as shown in Figure 1, was introduced by Zhang et al. [2] and has demonstrated strong performance in image denoising through residual learning and batch normalization. Instead of reconstructing clean images directly, DnCNN learns the noise component, enabling faster training and more accurate results.

The proposed architecture differs from the traditional DnCNN, which uses a sequence of convolutional layers to extract features, by introducing Internal Residual Learning (IRL) blocks to the architecture of the model. The use of IRL blocks allows incorporating identity skip connections in each IRL block, making it possible for the features to be propagated directly between layers, leading to better gradient flow during backpropagation and solving the problem of gradient disappearance faced by deep convolutional networks. This design improvement makes it possible for the network to learn more complex hierarchical representations while maintaining the structural details of CT images.

II. LITERATURE REVIEW

Image denoising has become a key research topic in medical image processing due to its importance in improving image quality and diagnostic accuracy.

Early strategies relied on spatial filters and transform-domain techniques. Gaussian filtering, despite its simplicity, tends to over-smooth images. Median filtering handles impulse noise well but may degrade fine structures. Wavelet-based methods operate in the frequency domain by reducing noisy coefficients using thresholding. [3]

Non-Local Means (NLM) [1] provided better denoising performance than earlier methods by taking advantage of the redundancy and similarity among image patches, helping preserve textures and edges. The BM3D [4] algorithm enhanced image denoising by combining block-matching techniques with collaborative filtering in a transformed domain, achieving superior performance compared to existing methods.

Recently, due to the advancements in deep learning techniques, there have been significant improvements in medical imaging denoising owing to the introduction of attention, transformers, and hybrid CNN-Transformer denoising models [9]. Such methods have shown better ability in capturing long-

Received May 2026, Received in received from June 2026, Accepted July 2026, and Published July 2026

range spatial dependencies and maintaining fine anatomical structures in low-dose CT scans. Furthermore, self-supervised and diffusion-based image denoising algorithms have gained popularity as viable methods in situations where high-quality reference images are not available [10]. However, most of the state-of-the-art methods are computationally costly and necessitate extensive computational resources for training the model [11]. Thus, lightweight denoising models with improved feature propagation, like the IRL-DnCNN model that has been suggested, can be considered effective solutions.

III. STRUCTURE OF CONVOLUTIONAL NEURAL NETWORK (CNN)

A Convolutional Neural Network (CNN) is a particular type of deep learning algorithm that takes input data in grid format, for instance, images or videos. The ability of CNNs to detect the presence of patterns like edges or textures makes CNNs the basis of computer vision [4]. The structure of a CNN consists of an input layer, an output layer, and multiple hidden layers
:

- **Convolution Layers**: This is the fundamental unit in a CNN. The filters (or kernels) are used to perform mathematical dot products across the entire input, thereby generating feature maps that are representative of certain features like corners or color contrasts.

- **Activation Functions**: Most CNN architectures use the ReLU (rectified linear unit) activation function, which comes after convolution. It converts all negative values into zeros; thus, making it non-linear so that the model learns complex non-linear patterns.
- **Pooling Layers**: This is commonly referred to as down sampling. The techniques used in pooling include Max-Pooling. It is a way to down-sample patches of a feature map by retaining the most significant signal values from them.
- **Fully Connected Layers:** At the end of the network, the processed features are flattened into a 1-dimensional vector and fed into a standard neural network. These layers perform high-level reasoning to produce the final classification or prediction.[5].

With the rise of deep learning, new architectures have outperformed classical methods. One of the best methods is DnCNN, which stands out due to :

- Residual learning framework.
- Batch normalization for stable and efficient training.
- Strong generalization across noise levels.
- High restoration accuracy.

These qualities make DnCNN suitable for medical imaging applications where detail preservation is critical.[2]

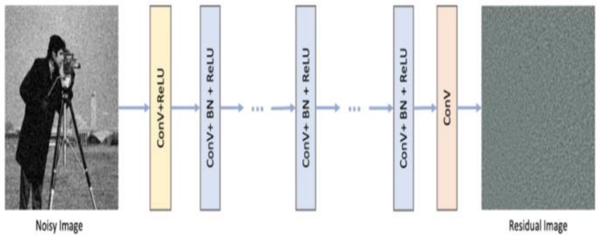

Figure 1. DnCNN architecture

IV. PROPOSED METHODOLOGY AND ARCHITECTURE DESCRIPTION

❖ ARCHITECTURE DESIGN: INTERNAL RESIDUAL-BASED DnCNN

The IRL-DnCNN model is developed in order to perform the residual noise estimation from CT images through the utilization of deep residual learning methodology. The network accepts as input the grayscale CT images with a dimensional size of 256 × 256 × 1. The feature extraction process starts with the first convolutional layer with 64 filters of size 3 × 3 with the same padding for the input, followed by the ReLU activation function. The main component of the suggested network is composed of nine internal residual learning blocks forming the network from a total of 20 layers (the first convolutional layer, 18 convolutional layers in the nine IRL blocks, and one final convolutional layer). Each IRL block contains two consecutive convolutional layers that utilize the 64 filters of size 3 × 3. After each convolution, the batch normalization is employed. The first batch normalization is followed by the ReLU activation function while the second batch normalization adds the produced output to the input via an identity skip connection element-wise. The application of this internal residual technique allows for easier propagation of the gradients during the backpropagation making the network immune to the vanishing gradient problem and leads to the possibility of the effective reuse of the features.
Finally, a 3 × 3 convolutional layer with a single output filter estimates the residual noise image. The denoised CT image is reconstructed by subtracting the predicted residual noise from the noisy input image, following the residual learning formulation originally introduced in DnCNN. Figure 2 illustrates the structure

of the Proposed Internal Residual Learning DnCNN (IRL_DnCNN).

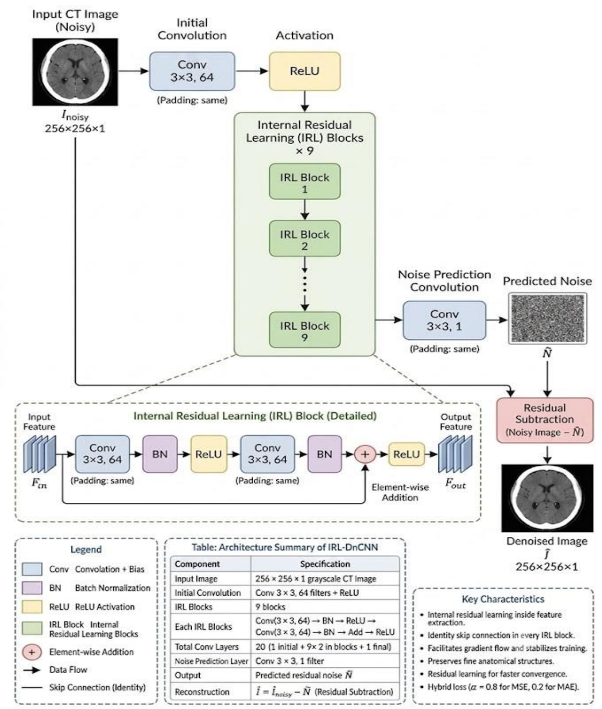

Figure 2 Proposed IRL-DnCNN Architecture

❖ LEARNING STRATEGY

Instead of learning a direct mapping from noisy images to clean images, DnCNN learns the residual noise components:

$$\mathbf{Y} = \mathbf{X} + \mathbf{V} \qquad (1)$$

Where:
X: clean image
V: noise
Y: noisy image

The network learns:

$$\mathbf{R}(\mathbf{Y}) \cong \mathbf{V} \qquad (2)$$

Then the denoised image is constructed as:

$$\hat{\mathbf{X}} = \mathbf{Y} - \mathbf{R}(\mathbf{Y}) \qquad (3)$$

This strategy of learning simplifies the optimization and accelerates the convergence. [2]

❖ OPTIMIZATION AND HYBRID LOSS FUNCTION

To achieve a balance between pixel-wise reconstruction accuracy and structural detail preservation, a hybrid loss function combining the Mean Squared Error (MSE) and Mean Absolute Error (MAE) is employed. The hybrid loss function is defined as:

$$L_{total} = \alpha \,.\, L_{MSE} + (1 - \alpha) \,.\, L_{MAE} \qquad (4)$$

The weighting coefficient was empirically selected as α=0.8 so that the MSE component was more weighted while maintaining the robustness of the MAE term. This design was chosen to fulfil two complementary objectives: minimising the overall reconstruction error and preserving fine anatomical structures. The dominant MSE term encourages accurate pixel-wise reconstruction and generally improves PSNR. The MAE term is less sensitive to outliers and helps to preserve edge information and reduce excessive image smoothing. Thus, the proposed hybrid loss provides a balanced optimisation objective for CT image denoising over different Gaussian noise levels

V. EXPERIMENTAL SETUP

❖ DATASET

The proposed IRL-DnCNN model was trained and evaluated using a publicly available brain CT image dataset obtained from the Kaggle platform [12]. The dataset consists of grayscale brain CT images collected for medical image analysis and diagnostic research. Prior to training, all images were resized to 256 × 256 pixels and normalized to the intensity range [0,1] to ensure numerical stability during optimization.

An additive Gaussian noise was synthetically introduced during data preparation. The standard deviation of the injected noise was randomly selected from the set σ ∈ {0.05, 0.10, 0.20, 0.30, 0.40}, enabling the proposed model to learn blind denoising across multiple noise levels. Rather than directly predicting the clean image, the network was trained to estimate the residual noise, which was obtained by subtracting the clean image from its noisy counterpart.

The original dataset provided for this study consists of individual CT image slices without explicit patient identifiers. Therefore, the dataset partitioning was performed at the image level. The complete dataset was randomly divided into 80% training, 10% validation, and 10% testing subsets. The training set was used for parameter optimization, the validation set for monitoring convergence and preventing overfitting, and the testing set for the final quantitative evaluation using PSNR and SSIM.

❖ NOISE GENERATION

Random Gaussian noise with levels ranging from 0.05 to 0.4 was added to the images during both training and testing stages. This allows the network to learn diverse noise characteristics and enhances its denoising capability, while evaluating its performance under challenging and varied noise conditions.

Gaussian noise is a kind of noise signal, the probability density function of which is equal to the normal distribution (also called Gaussian distribution). That is to say, the noise can only assume values which are normally distributed.

The probability density function p of a Gaussian random variable z is given by:

$$P(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \quad (5)$$

where z denotes the gray-level intensity, μ is the mean gray value, and σ is the standard deviation of the Gaussian distribution. This probability density function is used to model the Gaussian noise added to the CT images during the training process.[6]

❖ TRAINING CONFIGURATION

The model was trained using the following parameters:

- Network: IRL_DnCNN (20 layers)
- Image size: 256 * 256
- Batch size: 8.
- Epochs: 100.
- Learning rate: 1*10-3
- Reduce Learning rate on plateau factor = 0.5
- Patience = 5
- Optimizer: Adam.
- Loss function: Hybrid Loss (MSE and MAE).
- Programming language: Python.

❖ EVALUATION METRICS

The performance of the proposed method was conducted using:

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

PSNR evaluates the accuracy of reconstruction of the noisy image by comparing the denoised image to the clean image.

PSNR is most easily defined via the mean squared error (MSE). Given a noise-free $M * N$ image I and its noisy approximation K, MSE is defined as:

$$MSE = \frac{1}{M * N} \sum_{X=0}^{M-1} \sum_{Y=0}^{N-1} [F(X,Y) - I(X,Y)]^2 \quad (6)$$

The PSNR (in dB) is defined as:

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

$$PSNR = 20\log_{10}\left(\frac{MAXi}{\sqrt{MSE}}\right)$$

$$PSNR = 20\log_{10}(MAXi) - 10\log_{10}(MSE) \quad (7)$$

Where MAXi is the maximum possible pixel value of the image.[7]

- **Measure of Structural Similarity Index (SSIM)**

SSIM assesses the structural likeness and maintenance of visual image quality.

SSIM between two images x and y is defined as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2+\mu_y^2 + C_1)(\sigma_x^2+\sigma_y^2 + C_2)} \quad (8)$$

Where:

- μ_x , μ_y are the local mean intensities.
- σ_x^2, σ_y^2 are the local variances.
- σ_{xy} is the covariance between images x and y.
- C_1 and C_2 are stability constants.[8]

VI. EXPERIMENTAL RESULTS AND DISCUSSION

❖ EXPERIMENTAL RESULTS

All the experiments were performed using the TensorFlow/Keras deep learning framework on an NVIDIA Tesla P100 GPU to efficiently train and evaluate the proposed model. The model was tested on a set of testing images with noise levels between 0.05 and 0.4.

The results were compared to Gaussian filter, Non-local mean and standard DnCNN method at the same levels of noise as shown in table 1 below:

TABLE I. EXPERIMENTAL RESULTS

Noise Level	IRL-DnCNN	DnCNN	NLM	Gaussian filter
0.05	**PSNR= 36.0063 SSIM= 0.9603**	PSNR= 35.4401 SSIM= 0.9191	PSNR= 32.6632 SSIM= 0.4603	PSNR= 28.1312 SSIM= 0.4055
0.1	**PSNR= 33.1263 SSIM= 0.9308**	PSNR= 32.8912 SSIM= 0.9155	PSNR= 27.5903 SSIM= 0.3219	PSNR= 24.3427 SSIM= 0.3098
0.2	**PSNR= 30.1615 SSIM= 0.8951**	PSNR= 30.0320 SSIM= 0.8802	PSNR= 21.9488 SSIM= 0.2477	PSNR= 20.2698 SSIM= 0.2594
0.3	**PSNR= 28.2686 SSIM= 0.8719**	PSNR= 28.0913 SSIM= 0.8396	PSNR= 18.3830 SSIM= 0.2063	PSNR= 17.5239 SSIM= 0.2158
0.4	**PSNR= 26.7734 SSIM= 0.8491**	PSNR= 26.4142 SSIM= 0.7735	PSNR= 15.5605 SSIM= 0.1695	PSNR= 15.5665 SSIM= 0.1837

A sample of visual results of the proposed method and other methods is shown in figure 3.

A: is the original image.

B: represents the noisy images with different levels of noise (σ)

C: Denoised images using IRL_DnCNN.

D: Denoised images using DnCNN.
E: Denoised images using NLM.
F: Denoised images using a Gaussian filter.

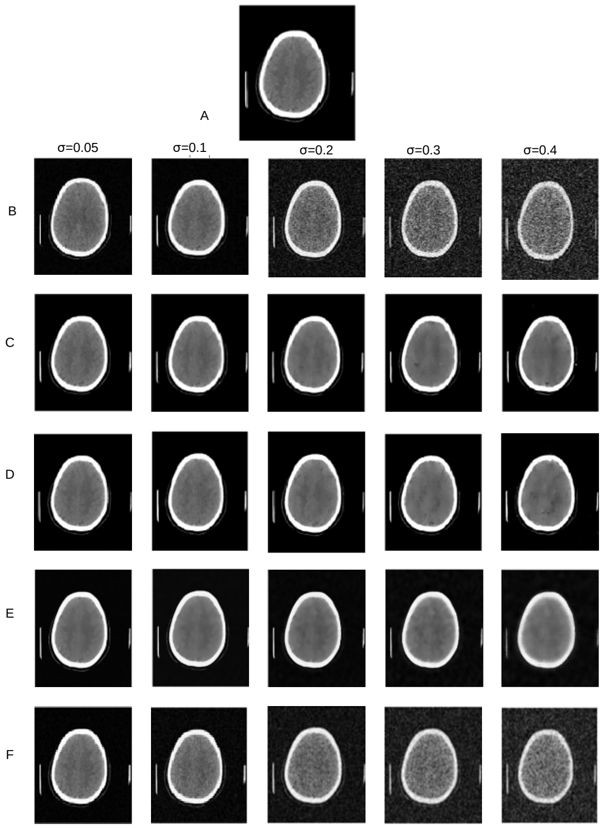

Figure 3: Sample of visual results

Figure 4 illustrates the loss curve of the proposed model. While figure 5 shows the loss curve of standard DnCNN. Finally figures 6 and 7, respectively, show the PSNR and SSIM values of the proposed method compared to the standard DnCNN, NLM, and Gaussian filter.

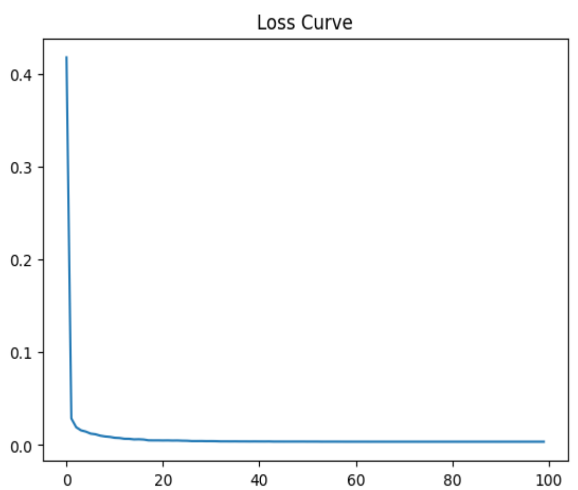

Figure 4: Loss Curve of the proposed model(IRL_DnCNN)

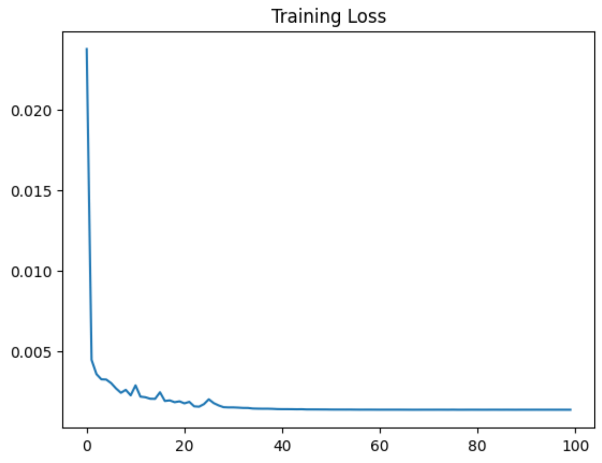

Figure 5: Loss Curve of standard DnCNN

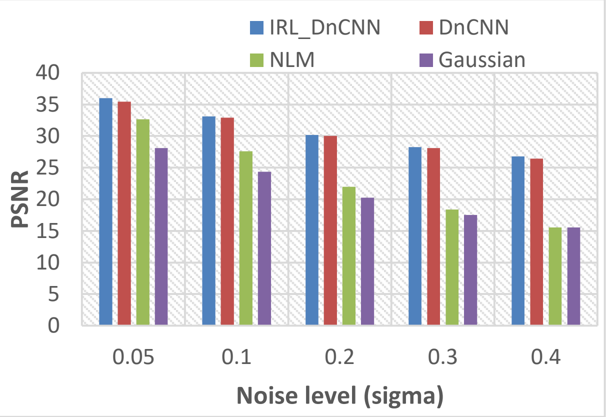

Figure 6: PSNR values.

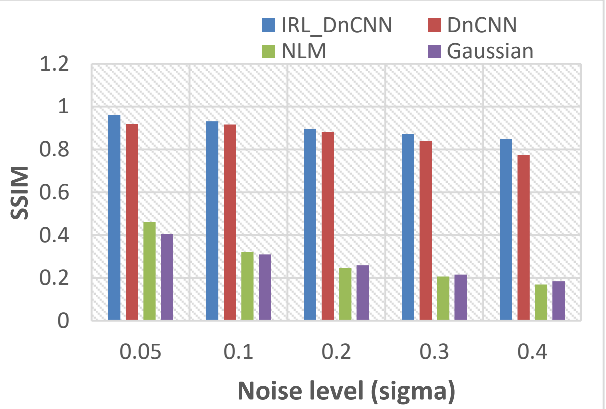

Figure 7: SSIM values.

❖ DISCUSSION

The experimental results demonstrate that the proposed IRL-DnCNN consistently achieves effective denoising performance across multiple Gaussian noise levels while preserving the structural integrity of CT images. The observed improvements can be attributed to several complementary design choices incorporated into the proposed architecture.

First, the integration of Internal Residual Learning (IRL) blocks facilitates efficient gradient propagation throughout the network. By employing identity skip connections within each residual block, low-level image features are directly propagated to deeper layers, reducing optimization difficulty and mitigating the vanishing-gradient problem. This enables the network to learn richer hierarchical feature representations while preserving fine anatomical structures that are essential for accurate medical diagnosis.

Second, the residual learning strategy simplifies the denoising task by requiring the network to estimate only the noise component rather than reconstructing the complete clean image. Since the residual noise generally exhibits lower structural complexity than the original image, the optimization process becomes more stable and converges more efficiently.

Furthermore, the proposed hybrid loss function combines the advantages of Mean Squared Error (MSE) and Mean Absolute Error (MAE). The dominant MSE component encourages accurate pixel-wise reconstruction, whereas the MAE component contributes to preserving edges and reducing excessive smoothing. This balanced optimization objective explains the simultaneous improvement in PSNR and SSIM observed during experimental evaluation.

The blind multi-noise training strategy further enhances the robustness of the proposed model. By exposing the network to multiple Gaussian noise levels during training, the learned representation becomes less dependent on a single degradation level and generalizes more effectively across varying noise conditions. Although the denoising performance gradually decreases as the noise level increases, the proposed network maintains stable reconstruction quality, indicating strong generalization capability.

From a clinical perspective, preserving subtle anatomical structures while effectively suppressing image noise is essential for improving diagnostic confidence in low-dose CT imaging. The proposed IRL-DnCNN achieves this objective through a computationally efficient architecture that balances denoising accuracy, structural preservation, and training stability.

VII. CONCLUSIONS

This study implemented an enhanced DnCNN-based framework for denoising CT images using a 20-layer convolutional network trained on 4427 Stroke-related CT images. The model leverages deep learning to effectively estimate and remove noise. Both quantitative and visual results confirm the model's effectiveness in reducing noise from CT images while maintaining structural integrity, compared to a Gaussian filter, Non-local mean, and standard DnCNN.

Future work for enhancing the performance of the model may include increasing training duration, expanding dataset size or validation on real low-dose clinical datasets, for further advanced practical applicability.

REFERENCES

1. A. Buades, B. Coll and J. M. Morel, "A non-local algorithm for image denoising". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2005, pp 60-65.
2. 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, 2017, pp. 3142-3155.
3. B. Goyal, A. Dogra, A. Agrawal, S. B. Sohi, and A. Shamra, "Image denoising review: From classical to state-of-the-art approaches", *Information Fusion*, vol. 55, 2020, pp. 220-244.
4. 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, 2007, pp. 2080-2095.
5. Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning", *Nature*, vol. 521, no. 7553, 2015, pp. 436-444.
6. P. Cattin, " image Restoration: Introduction to signal and Image Processing". *MAIC, University of Basel*, Apr. 24, 2012.
7. Q. Huynh-Thu and M. Ghanbari, " Scope of Validity of PSNR in Image/Video Quality Assessment", *Electronic Lectures*, vol. 44, no. 13, 2008, pp. 800-801.
8. Z. Wang, A. C. Bovik, H. R. Siekh and E. P. Simoncille, " Image quality assessment: From error visibility to structural similarity", *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004, pp. 600-612.
9. H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation," *Lecture Notes in Computer Science (MICCAI Workshops)*, vol. 12928, pp. 205–218, 2022.
10. J. Batson and L. Royer, "Noise2Self: Blind Denoising by Self-Supervision," *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pp. 524–533, 2019
11. J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. Yuille, and Y. Zhou, "TransUNet: Transformers Make Strong Encoders for Medical

Image Segmentation," *Medical Image Analysis*, vol. 85, Art. no. 102704, 2024.

12. U. Koç, E. Akçapınar Sezer, Y. Alper Özkaya, et al., "Artificial intelligence in healthcare competition (TEKNOFEST-2021): Stroke dataset," *Eurasian Journal of Medicine*, vol. 54, no. 3, pp. 248–258, 2022.