IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 October 2024, accepted 2 January 2025, date of publication 7 January 2025, date of current version 13 January 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3526619

RESEARCH ARTICLE

A Novel Attention-Guided Enhanced U-Net With Hybrid Edge-Preserving Structural Loss for Low-Dose CT Image Denoising

MUHAMMAD ZUBAIR[1], **HELMI MD RAIS**[1], **AND TALAL ALAZEMI**[2]

[1]Institute of Emerging Digital Technologies (EDiT), Center For Cyber Physical Systems (C2PS), Universiti Teknologi PETRONAS, Seri Iskandar 32160, Malaysia
[2]Department of Electronic Electrical Engineering, Brunel University London, UB8 3PH Uxbridge, U.K.

Corresponding author: Muhammad Zubair (muhammad_22000228@utp.edu.my)

This research was supported by the Institute of Emerging Digital Technologies (EDiT) & Center For Cyber Physical Systems (C2PS), Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia.

ABSTRACT Computed Tomography (CT) scan, pivotal for medical diagnostics, involves exposure to electromagnetic radiation, potentially elevating the risk of leukemia and cancer. Low-dose CT (LDCT) imaging has emerged to mitigate these risks, extensively reducing radiation exposure by up to 86%. However, it significantly reduces the quality of LDCT images and introduces noise and artifacts, degrading the diagnostic accuracy of the Computer Aided Diagnostic (CAD) system. This study presents a novel U-Net architecture, featuring several key enhancements. The model integrates residual blocks to improve feature representation and employs a custom hybrid loss function that combines structural loss with gradient regularization using the Euclidean norm, promoting superior CT image quality retention. Additionally, incorporating Attention Gates in the up-sampling layers of a proposed model optimizes the extraction of critical features, ensuring more precise denoising of CT images. The proposed model undergoes iterative training, using a custom loss function to refine its parameters and improve CT image denoising progressively. Its performance is rigorously evaluated both qualitatively and quantitatively on the '2016 Low-dose CT AAPM Grand Challenge dataset'. The results, assessed through the metrics Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Root Mean Square Error (RMSE), demonstrated promising improvements compared to state-of-the-art techniques. The model effectively reduces noise while preserving critical fine details, establishing itself as a highly efficient solution for LDCT image denoising.

INDEX TERMS Attention gate, deep learning, image enhancement, LDCT image denoising, residual blocks.

I. INTRODUCTION

In modern medicine, CT is a prevalent technique used to generate intricate cross-sectional images of the human body, enabling non-invasive detection of various pathological abnormalities like cancer, tumors, fractures, and vascular diseases at early stages [1], [2], [3], [4], [5], [6]. CT scans play a multifaceted role in advanced medical diagnostics [1], [3], [7]. However, ionizing radiation has diverse effects on the human body. Children exposed to the cumulative dose of three head scans are likely to develop leukemia and brain cancer, specifically, the risk of leukemia is significantly increased, with an odds ratio (OR) of 1.55 and a 95% confidence interval (CI) of 1.42 to 1.68 [3], [8]. Additionally, since the interior structure of veins is not visible in a regular CT scan, a supplementary substance known as *'iodinated contrast agent'* is used to boost the visibility of veins. In this situation, the radiation absorbed by the human body during a CT scan is doubled [1], [9], [10]. **FIGURE 1** compares radiation absorbed by the body parts during standard CT scans, when iodinated contrast agents are injected, and the duration required for the body to naturally absorb an equivalent dose of

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Afzal.

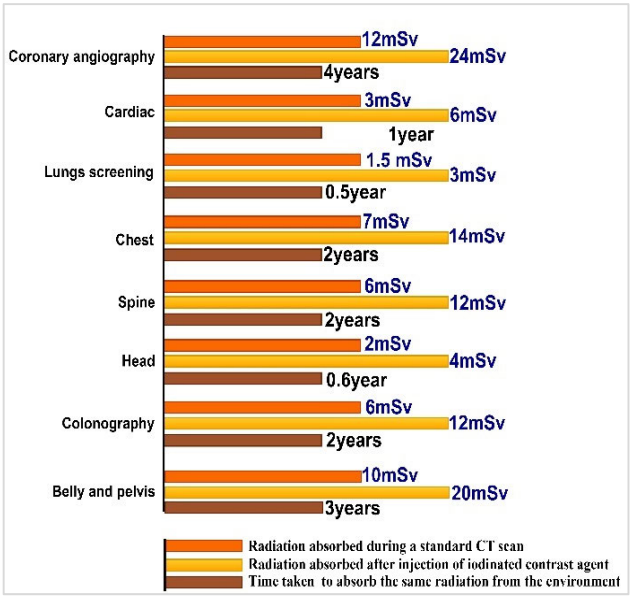

FIGURE 1. CT radiation absorption variability in the human body.

radiation from the environment, as represented by the diverse colors. Furthermore, to minimize the effect of ionized radiation, the ''As Low as Reasonably Achievable (ALARA)'' principle is adopted during the CT acquisition process [10], [11]. Additionally, X-ray flux is minimized to introduce an LDCT image to mitigate the health hazard. However, LDCT degrades the image quality and creates noise and artifacts, compromising the diagnostic accuracy of a CAD system [7], [12]. Consequently, various methods are used to uphold the image's quality and accuracy by alleviating the impact of these noise and artifacts [8], [13].

The LDCT image denoising techniques are broadly classified into three main categories, ''Preprocessing of Projection Data'', ''Statistical Iterative Reconstruction'' and ''Post-Reconstruction Image Processing'' [14], [15]. In preprocessing projection data techniques, any noise or artifacts present in the projection data are initially eliminated. Further, conventional reconstruction approaches such as Filtered Back Projection (FBP) are used to remove the noise and artifacts [16], [17]. However, Iterative Reconstruction (IR) methods surpass the traditional FBP and iteratively enhance CT image quality using synthetic raw data and prior information [18], [19]. Although IR methods significantly improve LDCT image quality, however, they increase computational complexity and pose a potential risk of content loss [18], [20].

The ''Post-Reconstruction Image Processing'' techniques are directly applied to constructed images and remove noise and artifacts [4], [21]. Further, these techniques are classified into two types to suppress the noise and artifacts in LDCT images i.e., ''Traditional methods'' and ''Deep learning-based methods'' [22], [23]. Traditional methods typically establish a straightforward connection between relevant information and image noise, after which optimization algorithms are employed to acquire denoised LDCT images [24]. Such methods usually rely on prior knowledge of noise [25]. Among traditional denoising techniques, ''Wavelet-based denoising'' and ''Block-matching 3D filter (BM3D)'' are popular ones [25], [26], [27].The former algorithms efficiently reduce noise by decomposing the image into different frequency scales. It excels in multi-resolution noise reduction while retaining image structures [7], [28]. Conversely, the later technique effectively preserves textures and minute details, offering robust denoising performance [4], [29]. Furthermore, ''Dictionary learning'' methods and ''Non-local Means (NLM)'' filters, both are advantageous techniques [3], [30], where ''Dictionary learning methods'' leverage data-driven dictionaries to capture complex patterns, making them effective for various noise types [23], [31]. Further, ''NLM filters'' exploit non-local similarities to preserve image structure, yielding impressive results even in cases of high noise levels [14], [32]. However, the balance between LDCT noise reduction and detail retention is still a complex and challenging task. To address these challenges, a novel framework has been proposed that effectively removes the noise and artifacts to enhance the quality of CT images. This approach aims to overcome the limitations of existing methods, ensuring clearer, more accurate CT image reconstructions. The main contribution of this paper is as follows.

1. An Attention Gate (AG) has been developed for LDCT image denoising, which selectively emphasizes significant features and regions while effectively filtering out noise and irrelevant details.
2. A hybrid loss function, termed SGLoss, has been proposed that combines the SSIM and Gradient loss with the Euclidean norm to improve perceptual quality and edge preservation in reconstructed CT images.
3. The proposed model, integrating Attention Gates (AG) and Structural Gradient Loss (SGLoss) within a U-Net framework enhanced by residual blocks, is evaluated on the 2016 Low-Dose CT AAPM Grand Challenge datasets using metrics such as PSNR, SSIM, and RMSE.
4. The results of the proposed model are compared with those of previous algorithms used for denoising LDCT images, demonstrating superior performance in both quantitative and visual quality. Additionally, an ablation study confirms that each component of the proposed model significantly enhances its overall effectiveness.

The structure of this paper is organized as follows: Section II provides a review of the relevant literature, while Section III delineates the methodology, offering a comprehensive overview of the proposed approach. Section IV presents an in-depth analysis of the results obtained, and Section V concludes the study with a summary of key findings and future work.

II. RELATED WORKS

Numerous efforts have been made to denoise LDCT images while preserving essential diagnostic information, utilizing

deep learning techniques. The discussion can be divided into three distinct categories, as outlined below.

A. DIVERSE DEEP LEARNING MODELS

Recent advancements in deep learning have shown significant promise in various fields, including medical imaging. Specifically, for LDCT noise reduction. Novel approaches like the three-layer Convolutional Neural Network (CNN) and the convolutional autoencoder-decoder with residual learning (RED-CNN) have emerged, offering effective solutions [33]. Additionally, deep CNNs incorporate directional wavelet transforms and shortcut connections for similar purposes in LDCT imaging [15], [33]. In the realm of LDCT image denoising, numerous advanced models with distinct techniques have emerged. However, overcoming persistent difficulties necessitates additional studies to enhance these models for the unique complexities of LDCT imaging. Likewise, Huang et al. [15] proposed TS-RCNN combining RCNN, wavelet transformations, and perceptual loss, though this is computationally complex. Furthermore, Chen et al. [16] proposed FRCNN using CNN, FTV loss, and residual learning for better noise reduction and detail preservation in LDCT images, however, it could be more computationally intensive than simpler denoising approaches. Additionally, Song et al. [17] proposed CMSNet, containing MSNet and a refinement network, predicts noise based on earlier layers of supervision, yet still, there are blurry effects. Similarly, Li et al. [18] proposed that ASWCNN uses wavelet transform and sub-pixel convolution in a top-down architecture. It maintains image structure and texture while removing noise and artifacts. However, further research is needed to enhance its performance. Furthermore, Kim et al. [20] introduced CNN denoisers for LDCT without real paired data, though imprecise kernel estimation can lead to errors in data synthesis. Additionally, Marcos et al. [21] utilized ResNet with dilated convolutions, batch normalization, and ReLU layers with fused spatial, and channel-attention modules, however, the proposed model is complex.

Furthermore, Gou et al. [22] introduce GRCNN by combining pixel-wise and gradient losses with the Sobel operator, nevertheless Sobel operator may not suit all structures and noise types. Moreover, Yan et al. [23] proposed a TLD-CDL framework, however, the suggested network may not fully capture the specific characteristics and complexities of LDCT images. Similarly, Shan et al. [24] proposed a MAP-NN consisting of a series of network modules, each incrementally improving image quality, though the model is complex. Additionally, Usui et al. [25] used DnCNN for denoising, including general CNN and transfer learning models. However, the method struggled with effective noise reduction at doses below 10% of the original, especially at incredibly low doses.

B. GENERATIVE ADVERSARIAL NETWORK(GAN)

GAN enhances the LDCT images through a generator network that transforms noisy images into cleaner versions, guided by a discriminator network for realism, resulting in improved image quality and diagnostic accuracy. Furthermore, Huang et al. [34] proposed DU-GAN, a novel approach utilizing GANs and U-Net-based discriminators to enhance LDCT image quality, however, the computational cost is higher. Likewise, Yi and Babyn [35] proposed SAGAN which incorporates adversarial and sharpness losses to address blur effects in LDCT images. Conversely, the proposed model detects subtle blurs, which could limit the improvement of overall sharpness. Moreover, Zhang et al. [36] introduced artifact and detail attention GAN, which uses a multi-channel generator to focus on noise, artifacts, edge features, and a multi-scale Res2Net discriminator to enhance discrimination. Yet, the computation cost is exceedingly high. Further, Ma et al. [37] utilize a noise learning Generative Adversarial Network (GAN) with specialized loss functions, including least squares, structural similarity, and L1 losses. However, it needs a large amount of data for effective training. Also, Wang and Hu [38] introduce a Progressive Wasserstein Generative Adversarial Network (PWGAN) with a Weighted Structurally Sensitive Hybrid loss Function (WSHL) to improve the LDCT image denoising. Further, Han et al. [26] proposed DESD, which contains a generator that includes a Pyramid Non-Local Attention (PNLA) module for feature correlation, and auxiliary shallow and deep feature processing modules to enhance feature extraction, However, the proposed network is sensitive to hyperparameters.

C. U-NeT ARCHITECTURE

This study utilized the deep encoder-decoder architecture, known as U-Net. The original U-Net was introduced by Ronneberger et al. [39] for medical image segmentation. The primary motivation for employing the U-Net framework in CT image denoising lies in its exceptional capability to extract hierarchical features across multiple spatial scales. This multi-scale approach enables U-Net to effectively identify and mitigate noise patterns that may vary in frequency and resolution, ensuring that critical structural details within the image are preserved. As a result, U-Net demonstrates a robust performance in enhancing image quality while maintaining the integrity of anatomical features, making it a preferred choice for medical imaging applications [39], [40]. Further, Zhang et al. [41] propose a CT image denoising method using U-Net and three attention modules: local, multi-feature channel, and hierarchical attention. The multi-attention mechanism effectively reduces noise while preserving structural details in CT images, however, it adds complexity. Also, Zubair et al. [13] introduced the U-Net autoencoder with dilated convolution and batch normalization layer to denoise the LDCT images, however, need to check on real LDCT datasets. Also, Zhang et al. [28] combine Transformer and CNN. The encoder CNN extracts local features, and the decoder includes a Dual-Path Transformer Block (DPTB) for structure and detail.

Multi-feature Spatial Attention Block (MSAB) enhances key regions and merely needs longer training. Moreover, Liu et al. [29] proposed ERA-WGAT which combines a residual autoencoder, edge enhancement, and a novel WGAT structure to address non-local information extraction in low-dose CT denoising, however utilizing much computational resources. Furthermore, Zubair et al. [42]. introduced DoG-UNet+, a model that integrates a DoG Sharpening Layer and an attention mechanism to enhance critical features, while the model effectively removes noise, its complexity is a drawback. Further, Jeon et al. [43] introduced MM-Net with a two-step training process involving MSAU-Net and U-Net-based denoiser. MM-Net is versatile and adaptable to various CT systems and protocols without ground-truth data. Still, it may exhibit lower denoising accuracy. Additionally, Liu et al. [44] proposed SureUnet, which addresses LDCT image artifacts and noise by integrating a sparse auto representation encoder with a U-Net architecture, yet, its practical feasibility in real-time has an impact.

III. METHODS

The methodology of the proposed model is further divided into the subparts below.

A. ATTENTION GATE

To denoise the LDCT image, an Attention Gate (AG) is introduced to enhance the feature selection in the U-Net architecture. The AG focuses on relevant regions of the feature maps and suppressing less important information, improving the denoising performance, especially around critical structures [45]. X is the input feature map, with a size of 255 × 255, as extracted from the LDCT image, after passing through several convolutional layers. It captures fine local details and textures of the image. While, G is the gating signal, which also has the same spatial dimensions, and is derived from a coarser feature map. It encapsulates more global context, providing a broader understanding of the image's structure and features. These two inputs are passed into the AG for further processing as given below.

1) *Linear Transformations:* Both the inputs X and G undergo linear transformation through convolution operations, denoted as $\phi_x(X)$ and $\phi_g(G)$ to reduce the dimensionality of the feature maps, making the subsequent operations more computationally efficient.
2) *Combination and non-linearity:* The transformed feature maps are then combined via element-wise addition $(\phi_x(X)+\phi_g(G)$ and passed through a non-linear activation function $\delta()$, such as ReLU. This step introduces non-linearity to the combined signal, enhancing its representational capacity.
3) *Linear Transformation:* The output from the non-linear activation is subjected to another linear transformation $\varphi()$, refining the combined feature representation and preparing it for the final attention mask computation.
4) *Sigmoid Normalization:* The refined signal is passed through a sigmoid function $\sigma()$, which normalizes the values to the range [0,1]. This produces the attention mask S, which highlights the regions of the input feature map X that are considered most important by the model.
5) *Attention Modulation:* The attention mask S is then applied to the original input feature map X of the LDCT image through element-wise multiplication $Y = S.X$ [46]. This operation ensures that only the important regions of X, as indicated by S, are emphasized in the output Y, the whole process has expressed in Equation (1).

$$S = \sigma(\phi(\delta(\phi_x(X) + \varphi_g(G)))) \tag{1}$$

The attention gates empower the U-Net model to concentrate on the most important features of the input LDCT images, facilitating more accurate reconstruction. Diagrammatically it can be represented as given in **FIGURE 2**.

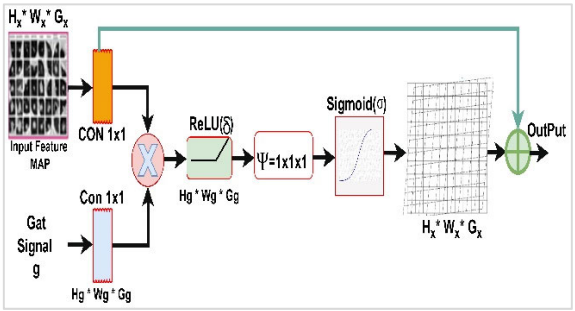

FIGURE 2. Attention Gate.

Where H represents the height, W stands for the width, and D indicates the depth, which corresponds to the number of channels of the CT image. In this study, H = 256, W = 256, and D = 1.

B. HYBRID LOSS

The objective of the proposed denoising method is to eliminate noise from CT images while preserving critical details, such as edges. However, minimizing the Mean Squared Error (MSE) between the denoised CT images and the NDCT images often leads to excessive blurring, which negatively impacts the perceived quality of the CT images. This approach is known to have a weak correlation with human assessments of CT image quality. Additionally, the ideal MSE estimator is prone to the regression-to-mean phenomenon, resulting in denoised LDCT images that may appear excessively smooth, unnatural, and unrealistic. In this study, a hybrid loss function is developed that incorporates the structure loss, SSIM, and Gradient loss to encapsulate both

structural similarity and edge preservation [47]. To assess the similarity between the denoised CT image and its normal-dose counterpart, the classic SSIM can be employed as given in Equation (2).

$$SSIM_{(x,y)} = \frac{(2u_x u_y + C_1) * (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) * (\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

where u_x, u_y, σ_x, σ_y and σ_{xy} represent the means, standard deviation, and correlation between the two images being compared. Additionally, C_1 and C_2 are the constants introduced to prevent division and eliminate the singularities. When x and y are nearly identical, the SSIM value approaches 1. In such cases, the SSIM loss can be expressed as shown in Equation (3).

$$Loss_{SSIM} = 1 - ssim\sum(x, y) \quad (3)$$

where SSIM refers to the Structural Similarity Index, a measure that ranges from -1 to 1, with 1 indicating identical images. When the SSIM value is close to 1, it suggests that the two images are highly similar, and as a result, the SSIM loss approaches zero. On the other hand, as the SSIM value decreases, the loss value increases. During the training the goal is to minimize this loss, which encourages the output image y to become increasingly like the reference CT image x. Gradient loss plays a critical role in preserving edge information and local details during the denoising process of LDCT images. Unlike traditional pixel-wise loss functions, which can obscure important structural features, gradient loss emphasizes the gradients in the image. This focus ensures that the model effectively retains vital details that are crucial for accurate interpretation, particularly in medical imaging scenarios. The horizontal and vertical gradients for both the ground truth CT image I_{true} and the predicted (denoised) CT image I_{pred} are calculated using Sobel kernels. Let G_u represent the *horizontal* Sobel kernel and G_v represent the *vertical* Sobel kernel, which are used to capture edge information in the respective directions. Further, the respective kernels are given in Equations (4) and (5).

$$Horizontal\ Sobel\ Kernel\ G_u = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (4)$$

$$Vertical\ Sobel\ Kernel\ G_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (5)$$

The horizontal and vertical gradients of both, the ground truth CT image I_{true} and the predicted CT image I_{pred} are computed by convolving them with Sobel kernel as given below. The Gradient of the ground truth CT image is given in Equation (6) and (7).

$$Horizontal\ Gradient = G_u(I_{true}) = I_{true} * G_u \quad (6)$$

$$Vertical\ Gradient = Gv(Itrue) = Itrue * Gv \quad (7)$$

Similarly, the Gradient for the predicted CT image is given below in Equations (8) and (9).

$$Horizontal\ Gradient = G_u(I_{pred}) = I_{pred} * G_u \quad (8)$$

$$Vertical\ Gradient = G_v(I_{pred}) = I_{pred} * G_v \quad (9)$$

For each pixel, the gradient magnitude using the Euclidean norm of the horizontal and vertical gradients at pixel P is calculated as given in Equation (10).

$$\|\nabla I_{true}(p)\| = \sqrt{G_u(i_{true}(p))^2 + G_v(i_{true}(p))^2} \quad (10)$$

where $G_u(i_{true})$ is the gradient in the x-direction (horizontal edges), and $G_v(i_{true})$ is the gradient in the y-direction (vertical edges). Similarly, the gradient magnitude of the predicted (i_{pred}) CT image at pixel P is given in given in Equation (11).

$$\left\|\nabla I_{pred}(p)\right\| = \sqrt{G_u(i_{pred}(p))^2 + G_v(i_{pred}(p))^2} \quad (11)$$

The gradient loss is computed by taking the squared difference between the gradient magnitudes of the ground truth and predicted images at each pixel, summing over all pixels, and normalizing by the total number of pixels N as given in Equation (12).

$$loss_{Grad} = \frac{1}{N}\sum_{P=1}^{N}(\|\nabla I_{true}(p)\| - \left\|\nabla I_{pred}(p)\right\|)^2 \quad (12)$$

Substituting the gradient magnitudes from Equations (10) and (11) as given in Equation (13).

$$loss_{Grad} = \frac{1}{N}\sum_{P=1}^{N}(G_u(i_{true}(p))^2 + G_v(i_{true}(p))^2 - (G_u(i_{pred}(p))^2 + G_v(i_{pred}(p))^2))^2 \quad (13)$$

This indicates the squared difference between the gradient magnitudes calculated using the Sobel filters in both horizontal and vertical directions of the ground truth CT image I_{true} and the predicted I_{pred}, summed over all pixels and normalized by the total number of pixels N [22]. So hybrid loss function is the combination of $^{Loss}{}_{SSIM}$ *and* L_{hybrid} as given in Equation (14).

$$L_{hybrid} = \alpha * 1 - ssim\sum(x, y) + \beta + \frac{1}{N}\sum_{P=1}^{N}(G_u(i_{true}(p))^2 + G_v(i_{true}(p))^2 - (G_u(i_{pred}(p))^2 + G_v(i_{pred}(p))^2))^2 \quad (14)$$

where the equation (14) represents a hybrid loss function L_{hybrid}, which combines SSIM and gradient-based loss to denoise the CT image. Further, x and y represent the reference (true) and predicted CT images, respectively. The SSIM loss focuses on the perceptual quality of the CT image by considering luminance, contrast, and structural features. This helps to reduce artifacts and blurring, maintaining the overall integrity of the CT image. Further, Gradient loss enhances the preservation of fine edges and transitions by comparing the gradients of the predicted and true images. Together, these loss functions provide a balanced approach to noise reduction, improving both the perceptual and clinical quality of LDCT images by minimizing noise and preserving essential features. Further, α and β are the coefficient or weighting

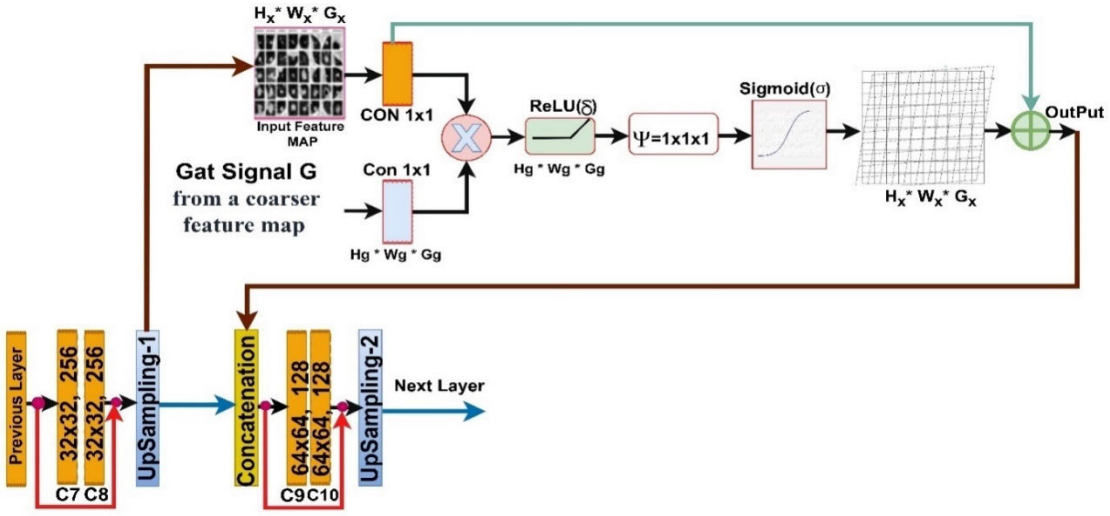

FIGURE 3. Residual blocks and Attention Gate.

factors that control the contribution of each individual loss component. Where parameter α controls the balance between SSIM and the gradient-based loss. A higher value of α gives greater importance to preserving the image's structure and details, which is critical for retaining diagnostic information in CT images. On the other hand, β emphasizes the preservation of edges details and fine structures, which are vital for accurate CT image denoising. In this study, based on trial and error, the values of $\alpha = 0.006$ and $\beta = 0.008$ were used.

C. NETWORK DESIGN

To address the limitations of basic U-Net architecture, this study enhances the U-Net model by introducing residual blocks, Attention gate, and hybrid loss function. The network begins with a $256 \times 256 \times 1$ input CT image that undergoes successive convolutional and max-pooling operations to reduce spatial dimensions while extracting increasingly complex features [48]. This down-sampling is complemented by residual blocks. The configuration of residual blocks with batch normalization following the convolution operations. This ensures that activations are normalized across the batch, maintaining a stable distribution of feature map values and mitigating issues like vanishing or exploding gradients, thus promoting smoother convergence during training. Mathematically the residual block can be represented as given in Equation (15).

$$Z = F(x) + X \tag{15}$$

Further, the network's deepest point, the bottleneck layer captures the abstract features representing the core structure of the CT image. The decoder then reconstructs the CT image by progressively sampling the features and concatenating them with corresponding layers from the encoder via skip connections, allowing the model to retain important spatial information. The concatenation layers ensure the dimensional consistency by aligning the spatial dimensions of feature maps from the encoder and decoder paths before merging them. This alignment is typically achieved through cropping the encoder feature maps to match the decoder dimensions, using padding during convolutions to prevent size mismatches. Similarly, the up-sampling operations in the decoder expand feature maps to the same size as their corresponding encoder layers. This process allows the U-Net to combine high-resolution spatial details from the encoder with contextual information from the decoder, preserving both low-level and high-level features for accurate image reconstruction. The addition of attention gates further enhances the focus on relevant features during up-sampling, improving the overall CT image quality [49]. The structure of the up-sampling layer in the proposed model, along with Residual blocks and Attention Gates is given in **FIGURE 3.**

The training of the enhanced proposed model utilizes a hybrid loss function that combines the SSIM and Gradient Loss to minimize reconstruction loss to enhance LDCT image quality. The SSIM component evaluates the similarity between reconstructed images and high-dose references by analyzing luminance, contrast, and structural information, ensuring that the generated CT images preserve essential visual characteristics critical for diagnosis. Meanwhile, Gradient Loss focuses on maintaining fine details and edges by assessing image gradients, thereby preventing blurring and ensuring sharpness in key regions. The optimization process iteratively adjusts model parameters to minimize a combined loss.

Through backpropagation, gradients of the total loss are computed, allowing for effective updates to the model weights. As training progresses and loss decreases, the proposed enhanced U-Net model generates high-quality CT images that are both statistically like high-dose references

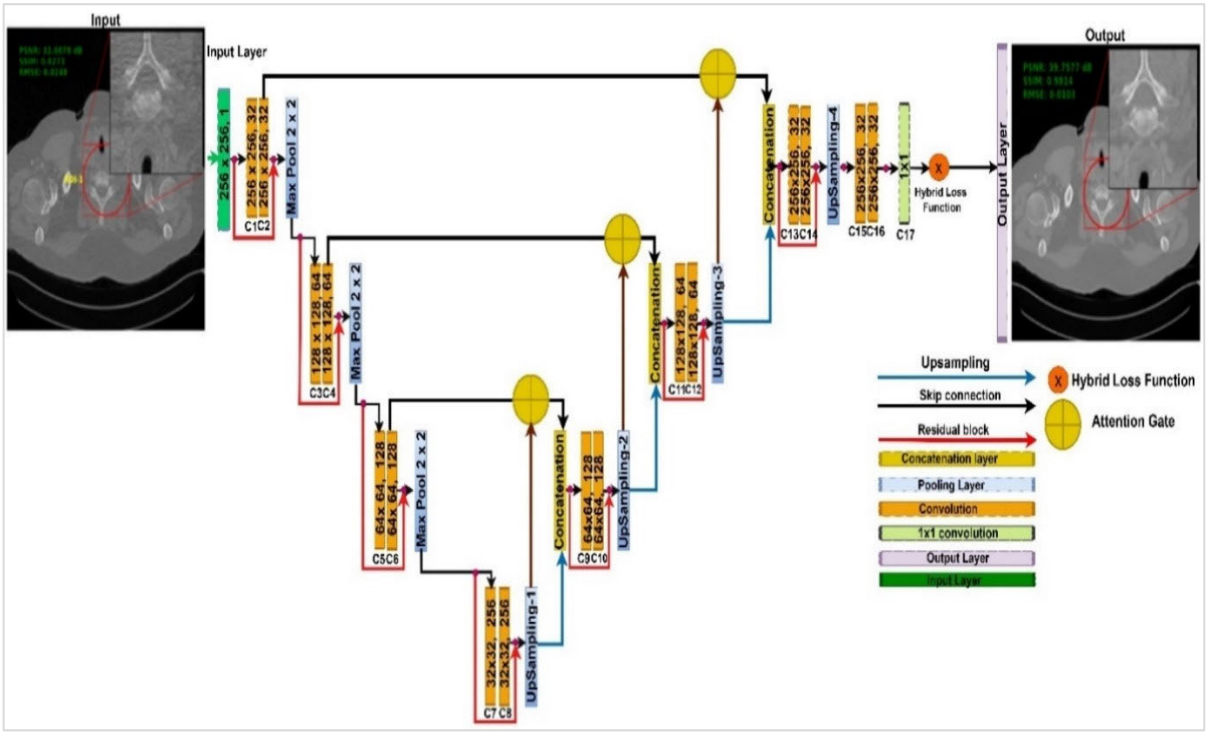

FIGURE 4. Proposed enhanced U-Net architecture with Residual Blocks, Attention Gate, and hybrid custom loss fusion.

and visually accurate, enhancing its capability for producing clinically relevant reconstructions [37]. The architecture of the enhanced proposed U-Net model is along the residual blocks, Attention Gate, and hybrid loss function is given in **FIGURE 4.**

IV. EXPERIMENT

A. DATASET FOR MODEL TRAINING

The enhanced U-Net model has been tested on the real '2016 Low-dose CT AAPM Grand Challenge Dataset,'which comprises 10 patient scans of the abdominal regions in DICOM format. These CT scans have a thickness of 1 mm and a dimension of 512 × 512 x 1. The dataset was partitioned into three parts: training, testing, and validation. The training set constituted 80% of the dataset, encompassing 4,617 pairs of LDCT images along with their corresponding high-dose counterparts. Further, the testing and validation subsets were meticulously assigned 10%, including 577 image pairs in each category. Each image in the dataset has been rescaled to the range [0,1]. To address overfitting and enhance the diversity and robustness of the improved U-Net model for better generalization to unseen data, two key augmentations are incorporated: horizontal flipping and random cropping. Additionally, early stopping is implemented to prevent overfitting by monitoring the validation loss during training and halting when performance on the validation set starts to deteriorate.

B. MODEL PARAMETERS

The proposed enhanced U-Net model utilizes a hybrid loss function that combines SSIM and gradient loss with weight parameters $\alpha = 0.006$ for SSIM and $\beta = 0.008$ for Gradient loss to effectively preserve structural similarity and fine-grained details in the denoised CT images. Further, the learning rate, set at 1.00E-03. The batch size is 32, indicating the number of samples processed before updating the model's weights. Training is conducted over 31 epochs, allowing the model to fully learn from the dataset. The Adam optimizer is used for efficient optimization. The ReLU activation function is applied to the output of each layer, ensuring non-linearity. A batch normalization is applied to normalize the inputs to each layer, promoting stable learning. The input images are 256 × 256 pixels with a single channel, corresponding to grayscale CT images.

C. PERFORMANCE COMPARISON

To assess the efficacy, a comparison was conducted between the proposed method and several classical techniques, including [50] BM3D, [51] NLM, [52] K-SVD, [53] CNCL, [30] WGAN-VGG, [12] EDCNN, and [54] CTformer. For quantitative evaluations, three assessment metrics are employed: PSNR, SSIM, and RMSE. The ablation study was conducted to evaluate the performance of the enhanced U-Net model by systematically assessing the impact of various components and modifications made to the architecture.

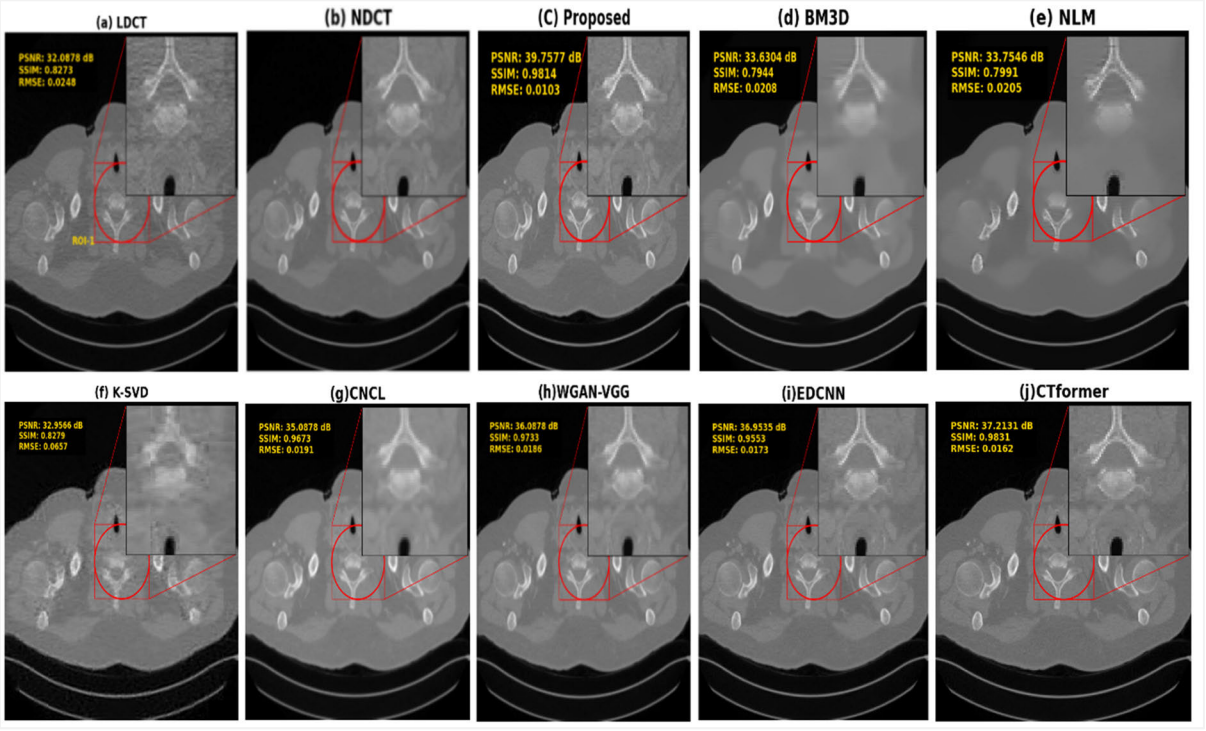

FIGURE 5. Comparison of denoising results of various methods for case-1.

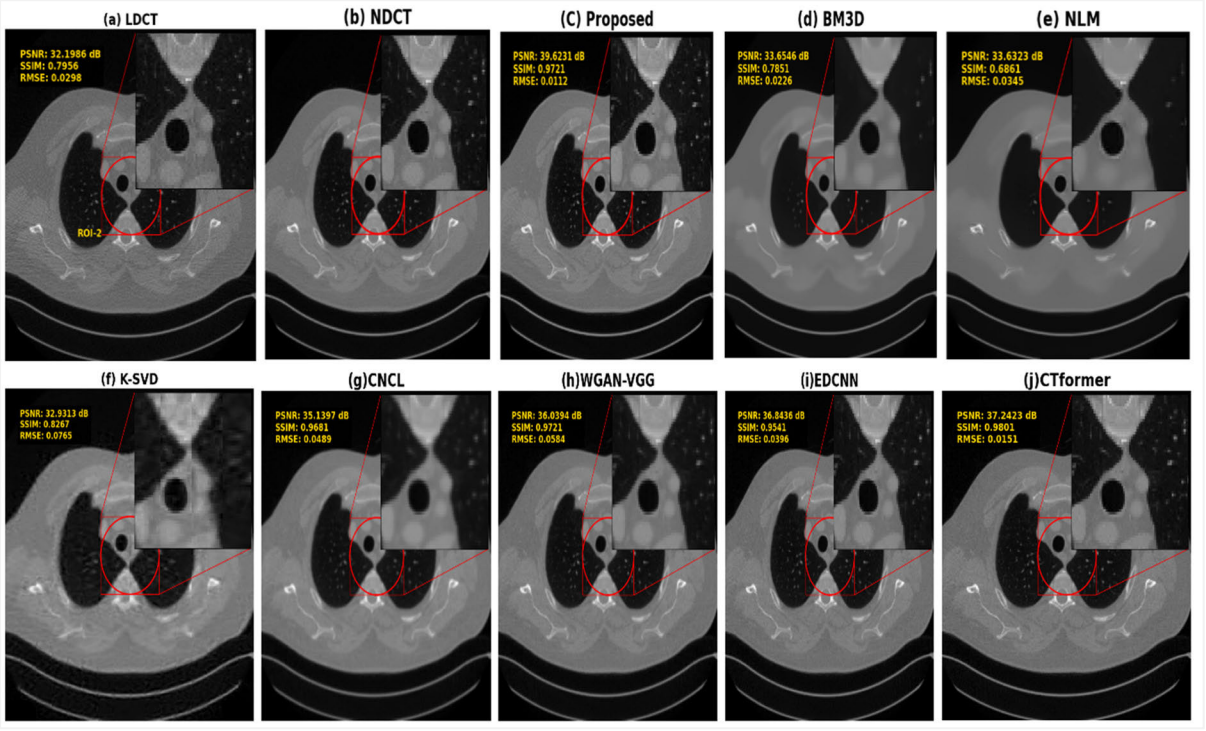

FIGURE 6. Comparison of denoising results of different of various methods for case-2.

D. MAYO EXPERIMENT

The denoising effects of various methods are illustrated on two representative slices, referred as Case-1 and Case-2, demonstrated in **FIGURE 5** and **FIGURE 6**. To illustrate the detailed preservation performance, a Region of Interest (ROI) has been selected in both cases and highlighted by a

TABLE 1. Quantitative results of the different methods (Means+SD).

S/N	Methods	PSNR	SSIM	RMSE
1	LDCT	32.1432 ± 0.0554	0.8114 ± 0.0158	0.0273 + 0.0025
2	BM3D	33.6425 ± 0.0121	0.7897 ± 0.0046	0.0217 + 0.0009
3	NLM	33.6935 ± 0.0611	0.7426 ± 0.0565	0.0275 + 0.0071
4	K-SVD	32.9439 ± 0.0126	0.8273 ± 0.0006	0.0711 + 0.0054
5	CNCL	35.1137 ± 0.0259	0.9677 ± 0.0004	0.0341 + 0.0149
6	WGAN-VGG	36.0636 ± 0.0242	0.9727 ± 0.0007	0.0385 + 0.0199
7	EDCNN	36.8985 ± 0.0549	0.9547 ± 0.0006	0.0284 + 0.0111
8	CTformer	37.2277 ± 0.0146	**0.9816 ± 0.0015**	0.0156 +0.0005
9	Proposed	**39.6904 + 0.0673**	0.9767 + 0.0046	**0.0107 + 0.0004**

red circle. For the *qualitative* analysis in Case-1, BM3D (d), NLM (e), and K-SVD (f) effectively reduce noise, but they blur fine anatomical structures. Furthermore, K-SVD introduces grainy textures along with artificial details, impacting the overall image quality. In Case-2, these traditional methods also suppress noise but perform poorly in preserving small structures, particularly in lung regions, where some fine details are either oversmoothed or lost.

In Case-2, CNCL (g) maintains a more balanced output compared to traditional approaches, providing better structural preservation. However, the finer textures are slightly smoothed out, affecting the clarity of smaller anatomical regions. WGAN-VGG (h) shows similar behaviours to Case-1, with effective noise reduction and structural retention, but in Case-2, the edges appear softer, and some areas lack sharpness. EDCNN (i) performs reasonably well, preserving key structures in both cases, but the residual noise is more pronounced in Case-2, affecting its overall visual quality.

Further, CTformer (j) demonstrates a good balance between noise suppression and structural integrity in both cases, yet it struggles slightly in Case-2 to maintain high contrast and detail sharpness, especially around complex lung textures. In contrast, the proposed method (c) demonstrates consistently superior performance in both Case-1 and Case-2, providing CT images that closely resemble the NDCT references CT images in terms of noise suppression, structural clarity, and contrast enhancement. In Case-1, the method effectively reduces noise while preserving fine anatomical structures, resulting in high PSNR and SSIM values. Similarly, in Case-2, it excels at maintaining intricate details and subtle textures without over-smoothing or introducing artifacts, achieving the best qualitative results among all the methods. This makes the proposed approach the most reliable solution across both cases, offering an optimal balance between noise reduction and structural preservation for low-dose CT image denoising. The visual comparison of the proposed methods along with classical methods is given in **FIGURE 5** and **FIGURE 6.**

Further, a *quantitative* comparison of various denoising methods applied to LDCT images, focusing on their performance in two distinct regions of interest (ROIs) as illustrated in **FIGURE 5** and **FIGURE 6**. The metrics utilized for comparison include PSNR, SSIM, and RMSE, with their corresponding average quantitative values, means, and standard deviations presented in **TABLE 1.** The highest values are highlighted in red, and the second-highest values are marked in sky blue.

Among traditional methods, BM3D and NLM perform reasonably well, with PSNR values around 33-34 dB and SSIM values below 0.8 in both cases. However, K-SVD exhibits the poorest performance, with high RMSE values and significantly lower PSNR and SSIM, reflecting excessive smoothing and loss of structural details. Additionally, the deep learning models demonstrate varying levels of denoising efficacy, with each exhibiting distinct strength. CNCL, the first of these, attains PSNR values of 35.08 dB in Case-1 and 35.13 dB in Case-2, alongside SSIM scores of 0.97 and 0.96, respectively. The results suggest there may be room for improvement compared to other advanced models. Furthermore, WGAN-VGG and EDCNN show competitive results, with PSNR values ranging from 36-37 dB and SSIM 97-95 scores. However, slightly higher RMSE values suggest a marginally reduced capacity for accurate noise suppression compared to CTformer. Also, CTformer clearly outperforms, achieving consistently high PSNR values of 37.21 dB in Case-1 and 37.24 dB in Case-2, alongside SSIM values exceeding 0.98. The notably low RMSE values in both cases, highlight CTformer's superior ability to maintain CT image quality and effectively suppress the noise. Finally, the proposed method outperforms all other CT denoising approaches in both PSNR and SSIM. However, CTformer achieves a slightly higher SSIM score of 0.0049 compared to the proposed method, highlighting its superior effectiveness in preserving structural details during the LDCT image denoising process.

The presented bar graphs in **FIGURE 7** provide a comprehensive comparison of various denoising techniques applied to LDCT images across two ROIs. The PSNR graph highlights the superior performance of the proposed model, achieving the highest peak signal-to-noise ratio, indicating enhanced image quality retention

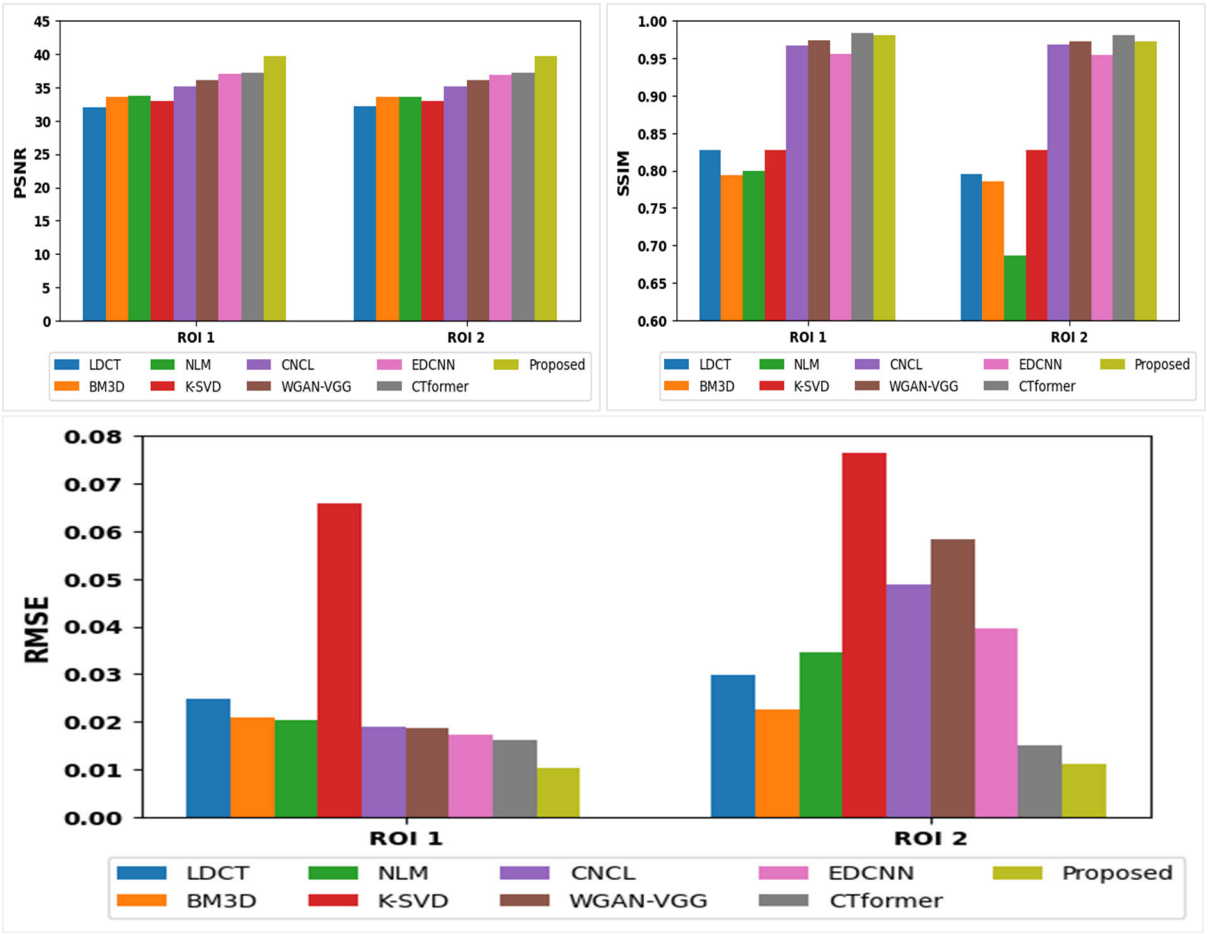

FIGURE 7. Quantitative performance of different ROIs in Figures 5 to 6.

TABLE 2. Quantitative performance of ablation studies (Means values).

Index	Base Model	MSE Loss	SSIM Loss	Gradient loss	Residual Blocks	Attention gate	PSNR	SSIM	RMSE
z	✔	✔	✕	✕	✕	✕	36.6545	0.9101	0.0166
B	✔	✕	✔	✕	✕	✕	37.3129	0.9354	0.0165
C	✔	✕	✔	✔	✕	✕	38.2147	0.9455	0.0135
D	✔	✕	✔	✔	✔	✕	39.5362	0.9561	0.0128
E	✔	✕	✔	✔	✔	✔	39.6381	0.9713	0.0123

in both ROI 1 and ROI 2. In the SSIM graph, the CTformar slightly shows better structural preservation, achieving high similarity to reference images, particularly in ROI 2.

The RMSE graph further emphasizes the model's effectiveness, showcasing the lowest root mean square error, which reflects minimal reconstruction errors across both regions. These graphs are essential for objectively validating the proposed model's ability to enhance image fidelity, preserve structural details, and outperform existing denoising approaches.

1) ABLATION STUDY

The ablation study assesses the impact of architectural components of any method. An ablation experiment was conducted to demonstrate the effectiveness of the proposed model using the ''2016 Low-dose CT AAPM Grand Challenge Dataset''. The average values of the selected ROIs from Case-3 in **FIGURE 8**, in terms of PSNR, SSIM, and RMSE, are provided in TABLE 2.

For simplicity, the proposed method is denoted as E, while the other variants of the proposed methods are labelled A, B, C, and D, respectively. A representative slice from Case-3

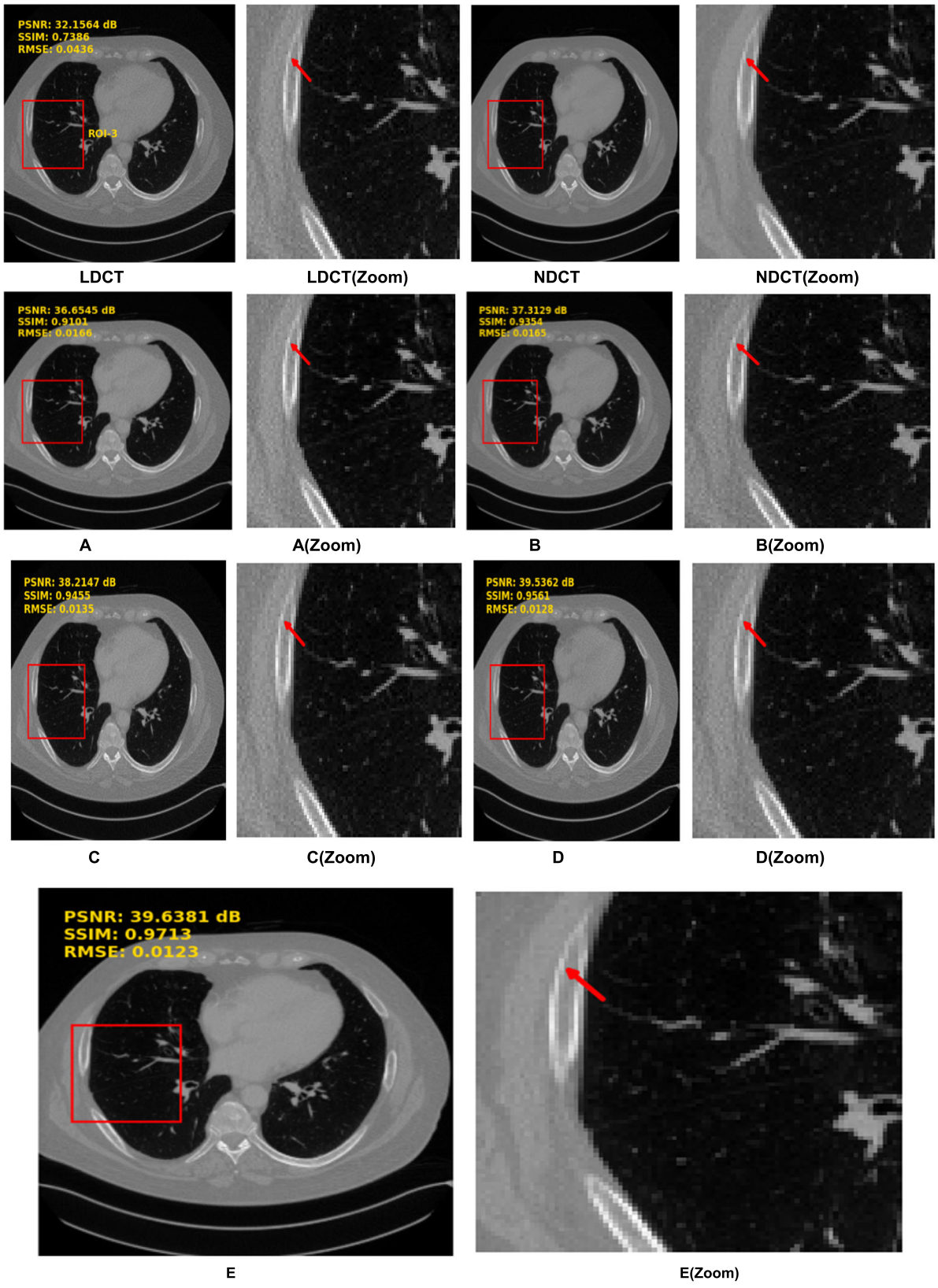

FIGURE 8. Comparison of denoising results of different of various methods for case-3.

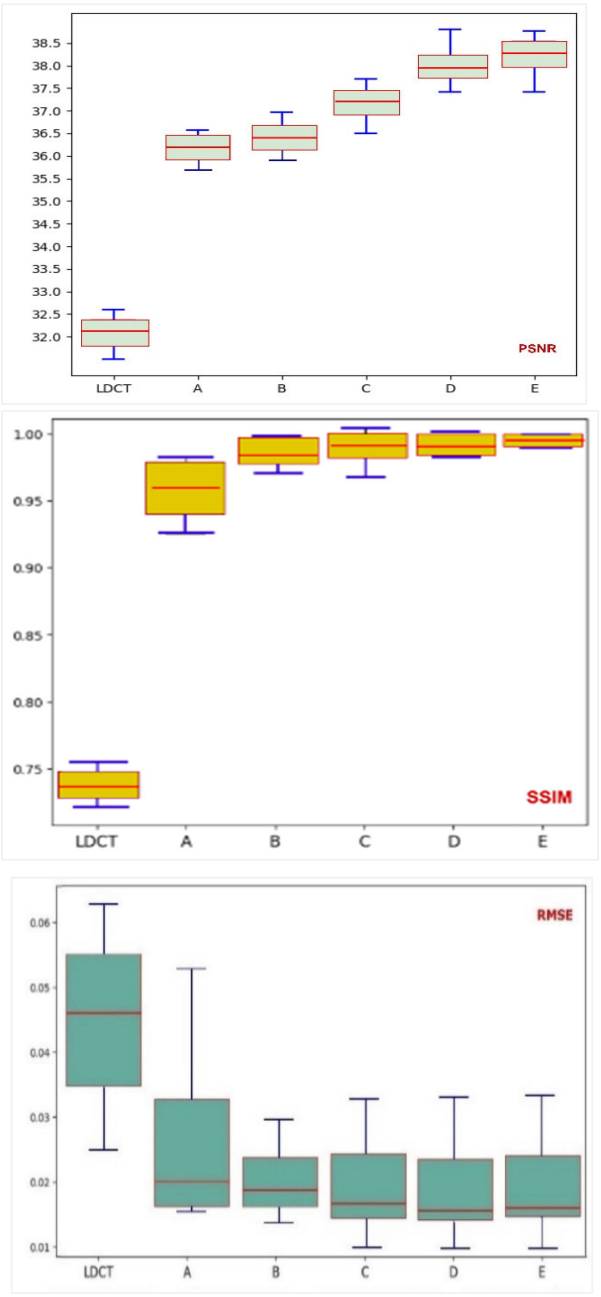

FIGURE 9. Boxplot of denoised results of ablation experiments on the mayo dataset.

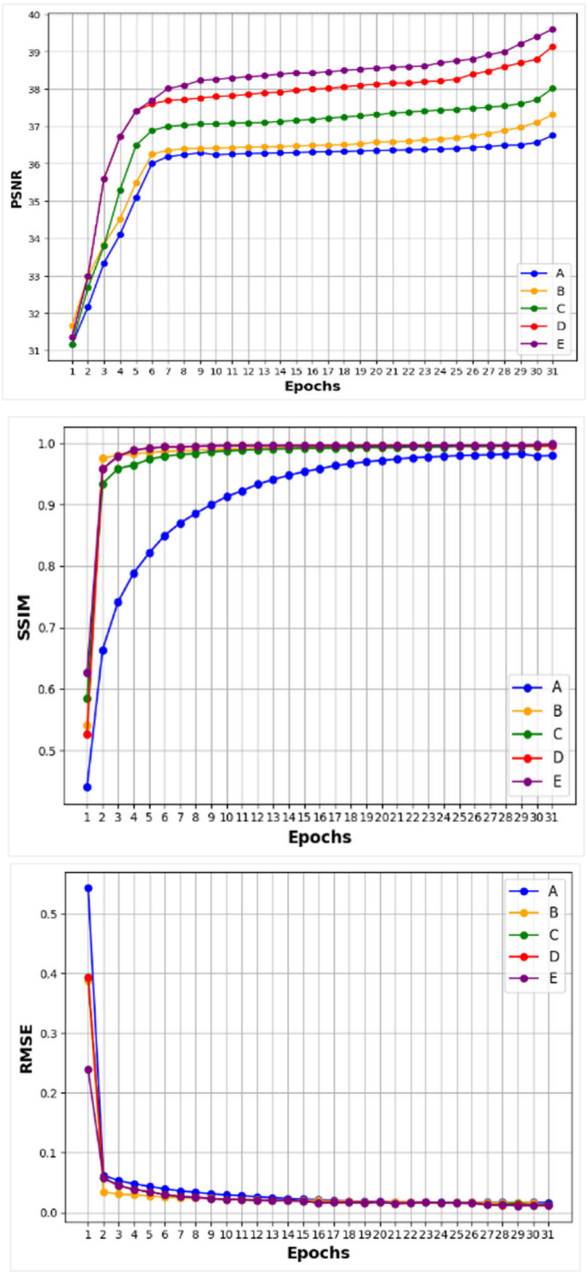

FIGURE 10. Line graph results of ablation experiments on the Mayo dataset.

was selected, focusing on ROI-3, with the highlighted regions marked by red rectangles.

The qualitative results for each proposed model variant are distinctly presented in the top-left corner of the respective denoised CT image as shown in **FIGURE 8**. where, Figure 8(A) and its zoomed version illustrate effective noise reduction; however, the preservation of structural details is suboptimal, leading to noticeable blurring in critical areas. Figure 8(B) and B(Zoom) demonstrate a significant improvement in structural preservation when compared to the MSE-based model, showcasing sharper anatomical boundaries and finer textural details. Nonetheless, a few minor noise artifacts remain detectable.

Further, Figure 8(C) presents enhanced denoising results relative to previous models, with marked improvements in image fidelity, perceptual quality, and average error. However, there remains room for enhancement in edge sharpness, texture preservation, and overall contrast. Furthermore, Figure 8(D), along with its zoomed counterpart D(zoom), clearly indicates a substantial improvement in overall quality and edge preservation compared to earlier model iterations. Nevertheless, some challenges regarding edge sharpness and contrast persist. Finally, Figure 8(E) and its zoomed

version further illustrate that the proposed model significantly surpasses prior versions in visual quality metrics. The enhancements in clarity, detail, edge preservation, and contrast yield a more visually appealing image, characterized by improved sharpness and reduced noise compared to earlier iterations.

The performance of different components of a proposed model used in the ablation study was evaluated on the test dataset by analysing PSNR, SSIM, and RMSE metrics through boxplots, as shown in **FIGURE 9.** These boxplots summarize the data using five key statistics: minimum, maximum, lower quartile, upper quartile, and median values. Based on the upper quartile values, model E demonstrates the highest robustness in terms of PSNR, SSIM, and RMSE. Additionally, the median values suggest that models D, and E lead in PSNR and SSIM, and RMSE.

Further, the line graphs of PSNR, SSIM, and RMSE over 31 epochs in **FIGURE 10** clearly demonstrate that Model D and E consistently outperform the other models in the ablation study, achieving higher PSNR and SSIM values, which indicate improved perceptual fidelity and reduced distortion. While the improvement in RMSE is more modest, However, Model E still demonstrates a clear advantage in minimizing pixel-wise errors, further reinforcing its robustness and overall performance in comparison to the other models.

V. DISCUSSION AND FUTURE WORK

This study presents an enhanced U-Net model designed for denoising LDCT images. The enhancements to the U-Net model are categorized into two main types: architectural improvements and the optimization of the cost function. The proposed cost function is a novel combination of structural similarity index (SSIM) and Gradient loss. The SSIM component focuses on preserving the perceptual quality of the CT images by assessing luminance, contrast, and structural elements, thereby effectively reducing artifacts and blurring while maintaining the structural integrity of CT images. Conversely, the gradient loss emphasizes the preservation of sharp edges and smooth transitions by analysing the gradient differences between the predicted and reference images. Together, these loss functions offer a comprehensive strategy for denoising, enhancing both the visual fidelity and clinical relevance of LDCT images by minimizing noise while protecting critical image features. The effectiveness of the proposed loss function is demonstrated through quantitative and qualitative results presented in the ablation study (TABLE 2, Index C) and illustrated in FIGURE 8(C), highlighting its contribution to performance improvements in the proposed model. Further, enhanced the architecture, deep residual blocks have been integrated, which facilitate improved gradient flow and mitigate the vanishing gradient problem, allowing the model to learn residuals more effectively. This approach not only enhances generalization and increases network depth but also preserves essential features within the images. The quantitative benefits of the residual blocks are evidenced in the ablation study (TABLE 2, Index D), while qualitative results are depicted in FIGURE 8(D). Furthermore, the incorporation of Attention Gates within the proposed model architecture has resulted in significant enhancements in both quantitative and qualitative metrics. This is corroborated by the results showcased in (TABLE 2, Index E) and FIGURE 8(E). The study presents an enhanced U-Net model for denoising low-dose CT (LDCT) images, featuring a novel cost function that combines structural similarity index (SSIM) and gradient loss with the Euclidean norm to effectively reduce noise while preserving critical image features. Additionally, incorporating deep residual blocks and Attention Gates enhances gradient flow and improves feature selection capabilities. Both qualitative and quantitative comparisons, as well as an ablation study, demonstrate that the proposed model outperforms benchmark algorithms in terms of PSNR and RMSE. Although the CTformer shows slightly better SSIM results in ROI-2 compared to other techniques, the overall performance of the proposed model validates its effectiveness in enhancing the quality of low-dose CT images. Future work will focus on validating and evaluating this technique on larger datasets to further enhance its generalizability and robustness.

DATA AVAILABILITY STATEMENT

The data supporting this study are available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declared that they had no conflict of interest.

ACKNOWLEDGMENT

The authors express their heartfelt gratitude to Universiti Teknologi PETRONAS, Malaysia, for providing the facilities essential for conducting this research. They also deeply appreciate the invaluable contributions of Mr. Ahmad Abubakar Suleiman and Mr. Arsalaan Khan Yousafzai from Universiti Teknologi PETRONAS, whose assistance greatly enriched the writing of this manuscript. Lastly, they are sincerely thankful to the anonymous editors and peer reviewers for their meticulous revisions and constructive feedback.

REFERENCES

[1] S. Dodia, B. Annappa, and P. A. Mahesh, "Recent advancements in deep learning based lung cancer detection: A systematic review," *Eng. Appl. Artif. Intell.*, vol. 116, Nov. 2022, Art. no. 105490, doi: 10.1016/j.engappai.2022.105490.

[2] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, and P. H. Abreu, "Using deep learning techniques in medical imaging: A systematic review of applications on CT and PET," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4093–4160, Aug. 2020, doi: 10.1007/s10462-019-09788-3.

[3] V. Sorin, Y. Barash, E. Konen, and E. Klang, "Creating artificial images for radiology applications using generative adversarial networks (GANs)—A systematic review," *Academic Radiol.*, vol. 27, no. 8, pp. 1175–1185, Aug. 2020, doi: 10.1016/j.acra.2019.12.024.

[4] K. A. S. H. Kulathilake, N. A. Abdullah, A. Q. M. Sabri, and K. W. Lai, "A review on deep learning approaches for low-dose computed tomography restoration," *Complex Intell. Syst.*, vol. 9, pp. 2713–2745, May 2021, doi: 10.1007/s40747-021-00405-x.

[5] Y.-H. Shao, K. Tsai, S. Kim, Y.-J. Wu, and K. Demissie, "Exposure to tomographic scans and cancer risks," *JNCI Cancer Spectr.*, vol. 4, no. 1, pp. 1–7, Feb. 2020, doi: 10.1093/jncics/pkz072.

[6] C. H. Schultz, R. Fairley, L. S.-L. Murphy, and M. Doss, "The risk of cancer from CT scans and other sources of low-dose radiation: A critical appraisal of methodologic quality," *Prehospital Disaster Med.*, vol. 35, no. 1, pp. 3–16, Feb. 2020, doi: 10.1017/s1049023x1900520x.

[7] Z. Zhang and E. Seeram, "The use of artificial intelligence in computed tomography image reconstruction—A literature review," *J. Med. Imag. Radiat. Sci.*, vol. 51, no. 4, pp. 671–677, Dec. 2020, doi: 10.1016/j.jmir.2020.09.001.

[8] M. Zhang, S. Gu, and Y. Shi, "The use of deep learning methods in low-dose computed tomography image reconstruction: A systematic review," *Complex Intell. Syst.*, vol. 8, no. 6, pp. 5545–5561, Dec. 2022, doi: 10.1007/s40747-022-00724-7.

[9] S. Bera and P. K. Biswas, "Noise conscious training of non local neural network powered by self attentive spectral normalized Markovian patch GAN for low dose CT denoising," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3663–3673, Dec. 2021. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9474492

[10] S. V. Mohd Sagheer and S. N. George, "A review on medical image denoising algorithms," *Biomed. Signal Process. Control*, vol. 61, Aug. 2020, Art. no. 102036, doi: 10.1016/j.bspc.2020.102036.

[11] C. Tan, M. Yang, Z. You, H. Chen, and Y. Zhang, "A selective kernel-based cycle-consistent generative adversarial network for unpaired low-dose CT denoising," *Precis. Clin. Med.*, vol. 5, no. 2, Apr. 2022, Art. no. pbac011, doi: 10.1093/pcmedi/pbac011.

[12] T. Liang, Y. Jin, Y. Li, and T. Wang, "EDCNN: Edge enhancement-based densely connected network with compound loss for low-dose CT denoising," in *Proc. 15th IEEE Int. Conf. Signal Process. (ICSP)*, vol. 1, Dec. 2020, pp. 193–198, doi: 10.1109/ICSP48669.2020.9320928.

[13] M. Zubair, H. B. M. Rais, and Q. Al-Tashi, "U-Net autoencoder for edge-preserved denoising of low dose computed tomography images: A novel technique," in *Proc. 13th Int. Conf. Inf. Technol. Asia (CITA)*, Aug. 2023, pp. 19–24, doi: 10.1109/CITA58204.2023.10262803.

[14] B. Gajera, S. R. Kapil, D. Ziaei, J. Mangalagiri, E. Siegel, and D. Chapman, "CT-scan denoising using a charbonnier loss generative adversarial network," *IEEE Access*, vol. 9, pp. 84093–84109, 2021, doi: 10.1109/ACCESS.2021.3087424.

[15] L. Huang, H. Jiang, S. Li, Z. Bai, and J. Zhang, "Two stage residual CNN for texture denoising and structure enhancement on low dose CT image," *Comput. Methods Programs Biomed.*, vol. 184, Feb. 2020, Art. no. 105115, doi: 10.1016/j.cmpb.2019.105115.

[16] M. Chen, Y.-F. Pu, and Y.-C. Bai, "Low-dose CT image denoising using residual convolutional network with fractional TV loss," *Neurocomputing*, vol. 452, pp. 510–520, Sep. 2021, doi: 10.1016/j.neucom.2020.10.004.

[17] H. Song, L. Chen, Y. Cui, Q. Li, Q. Wang, J. Fan, J. Yang, and L. Zhang, "Denoising of MR and CT images using cascaded multi-supervision convolutional neural networks with progressive training," *Neurocomputing*, vol. 469, pp. 354–365, Apr. 2022, doi: 10.1016/j.neucom.2020.10.118.

[18] S. Li, Q. Li, R. Li, W. Wu, J. Zhao, Y. Qiang, and Y. Tian, "An adaptive self-guided wavelet convolutional neural network with compound loss for low-dose CT denoising," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103543, doi: 10.1016/j.bspc.2022.103543.

[19] M. Zubair, H. Rais, Q. Al-tashi, and F. Ullah, "Enabling predication of the deep learning algorithms for low-dose CT scan image denoising models? A systematic literature review," *IEEE Access*, vol. 12, pp. 79025–79050, 2024, doi: 10.1109/ACCESS.2024.3407774.

[20] B. Kim, S. E. Divel, N. J. Pelc, and J. Baek, "A methodology to train a convolutional neural network-based low-dose CT denoiser with an accurate image domain noise insertion technique," *IEEE Access*, vol. 10, pp. 86395–86407, 2022, doi: 10.1109/ACCESS.2022.3198948.

[21] L. Marcos, J. Alirezaie, and P. Babyn, "Low dose CT denoising by ResNet with fused attention modules and integrated loss functions," *Frontiers Signal Process.*, vol. 1, pp. 1–11, Feb. 2022, doi: 10.3389/frsip.2021.812193.

[22] S. Gou, W. Liu, C. Jiao, H. Liu, Y. Gu, X. Zhang, J. Lee, and L. Jiao, "Gradient regularized convolutional neural networks for low-dose CT image enhancement," *Phys. Med. Biol.*, vol. 64, no. 16, Aug. 2019, Art. no. 165017, doi: 10.1088/1361-6560/ab325e.

[23] R. Yan, Y. Liu, Y. Liu, L. Wang, R. Zhao, Y. Bai, and Z. Gui, "Image denoising for low-dose CT via convolutional dictionary learning and neural network," *IEEE Trans. Comput. Imag.*, vol. 9, pp. 83–93, 2023, doi: 10.1109/TCI.2023.3241546.

[24] H. Shan, A. Padole, F. Homayounieh, U. Kruger, R. D. Khera, C. Nitiwarangkul, M. K. Kalra, and G. Wang, "Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction," *Nature Mach. Intell.*, vol. 1, no. 6, pp. 269–276, Jun. 2019, doi: 10.1038/s42256-019-0057-9.

[25] K. Usui, K. Ogawa, M. Goto, Y. Sakano, S. Kyougoku, and H. Daida, "Quantitative evaluation of deep convolutional neural network-based image denoising for low-dose computed tomography," *Vis. Comput. Ind., Biomed., Art*, vol. 4, no. 1, p. 21, Dec. 2021, doi: 10.1186/s42492-021-00087-9.

[26] Z. Han, H. Shangguan, X. Zhang, P. Zhang, X. Cui, and H. Ren, "A dual-encoder-single-decoder based low-dose CT denoising network," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 7, pp. 3251–3260, Jul. 2022, doi: 10.1109/JBHI.2022.3155788.

[27] G. L. Zeng, "An attempt of directly filtering the sparse-view CT images by BM3D," *Proc. SPIE*, vol. 12304, no. 1, pp. 568–575, Oct. 2022, doi: 10.1117/12.2646426.

[28] J. Zhang, Z. Shangguan, W. Gong, and Y. Cheng, "A novel denoising method for low-dose CT images based on transformer and CNN," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107162, doi: 10.1016/j.compbiomed.2023.107162.

[29] H. Liu, P. Liao, H. Chen, and Y. Zhang, "ERA-WGAT: Edge-enhanced residual autoencoder with a window-based graph attention convolutional network for low-dose CT denoising," *Biomed. Opt. Exp.*, vol. 13, no. 11, p. 5775, 2022, doi: 10.1364/boe.471340.

[30] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1348–1357, Jun. 2018, doi: 10.1109/TMI.2018.2827462.

[31] Y. Chen, L. Shi, Q. Feng, J. Yang, H. Shu, L. Luo, J.-L. Coatrieux, and W. Chen, "Artifact suppressed dictionary learning for low-dose CT image processing," *IEEE Trans. Med. Imag.*, vol. 33, no. 12, pp. 2271–2292, Dec. 2014, doi: 10.1109/TMI.2014.2336860.

[32] M. Diwakar, P. Singh, C. Swarup, E. Bajal, M. Jindal, V. Ravi, K. U. Singh, and T. Singh, "Noise suppression and edge preservation for low-dose COVID-19 CT images using NLM and method noise thresholding in shearlet domain," *Diagnostics*, vol. 12, no. 11, p. 2766, Nov. 2022, doi: 10.3390/diagnostics12112766.

[33] M. M. Lell and M. Kachelrieß, "Recent and upcoming technological developments in computed tomography: High speed, low dose, deep learning, multienergy," *Investigative Radiol.*, vol. 55, no. 1, pp. 8–19, 2020, doi: 10.1097/rli.0000000000000601.

[34] Z. Huang, J. Zhang, Y. Zhang, and H. Shan, "DU-GAN: Generative adversarial networks with dual-domain U-net-based discriminators for low-dose CT denoising," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022, doi: 10.1109/TIM.2021.3128703.

[35] X. Yi and P. Babyn, "Sharpness-aware low-dose CT denoising using conditional generative adversarial network," *J. Digit. Imag.*, vol. 31, no. 5, pp. 655–669, Oct. 2018.

[36] X. Zhang, Z. Han, H. Shangguan, X. Han, X. Cui, and A. Wang, "Artifact and detail attention generative adversarial networks for low-dose CT denoising," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3901–3918, Dec. 2021, doi: 10.1109/TMI.2021.3101616.

[37] Y. Ma, B. Wei, P. Feng, P. He, X. Guo, and G. Wang, "Low-dose CT image denoising using a generative adversarial network with a hybrid loss function for noise learning," *IEEE Access*, vol. 8, pp. 67519–67529, 2020, doi: 10.1109/ACCESS.2020.2986388.

[38] G. Wang and X. Hu, "Low-dose CT denoising using a progressive Wasserstein generative adversarial network," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104625, doi: 10.1016/j.compbiomed.2021.104625.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Jan. 2015, pp. 234–241.

[40] W. Jifara, F. Jiang, S. Rho, M. Cheng, and S. Liu, "Medical image denoising using convolutional neural network: A residual learning approach," *J. Supercomput.*, vol. 75, no. 2, pp. 704–718, Feb. 2019, doi: 10.1007/s11227-017-2080-0.

[41] J. Zhang, Y. Niu, Z. Shangguan, W. Gong, and Y. Cheng, "A novel denoising method for CT images based on U-net and multi-attention," *Comput. Biol. Med.*, vol. 152, Jan. 2023, Art. no. 106387, doi: 10.1016/j.compbiomed.2022.106387.

[42] M. Zubair, H. B. M. Rais, F. Ullah, A. K. Yousafzai, and F. Hassan, "Enhancing low-dose CT image quality through deep learning: A DoG-sharpened U-Net approach with attention mechanism," in *Proc. ASU Int. Conf. Emerg. Technol. Sustainability Intell. Syst. (ICETSIS)*, Jan. 2024, pp. 1037–1041, doi: 10.1109/ICETSIS61505.2024.10459568.

[43] S.-Y. Jeon, W. Kim, and J.-H. Choi, "MM-Net: Multiframe and multimask-based unsupervised deep denoising for low-dose computed tomography," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 7, no. 3, pp. 296–306, Mar. 2023, doi: 10.1109/TRPMS.2022.3224553.

[44] J. Liu, T. Zhang, Y. Kang, J. Qiang, D. Hu, and Y. Zhang, "SureUnet: Sparse autorepresentation encoder U-Net for noise artifact suppression in low-dose CT," *Neural Comput. Appl.*, vol. 2023, pp. 1–12, Jul. 2023, doi: 10.1007/s00521-023-08847-9.

[45] S. Ding, Q. Wang, L. Guo, J. Zhang, and L. Ding, "A novel image denoising algorithm combining attention mechanism and residual UNet network," *Knowl. Inf. Syst.*, vol. 66, no. 1, pp. 581–611, Jan. 2024, doi: 10.1007/s10115-023-01965-9.

[46] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, and D. Rueckert, "Attention gated networks: Learning to leverage salient regions in medical images," *Med. Image Anal.*, vol. 53, pp. 197–207, Apr. 2019, doi: 10.1016/j.media.2019.01.012.

[47] Z. Li, W. Shi, Q. Xing, Y. Miao, W. He, H. Yang, and Z. Jiang, "Low-dose CT image denoising with improving WGAN and hybrid loss function," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–14, Aug. 2021, doi: 10.1155/2021/2973108.

[48] Z. Li, J. Huang, L. Yu, Y. Chi, and M. Jin, "Low-dose CT image denoising using cycle-consistent adversarial networks," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSS/MIC)*, Oct. 2019, pp. 1–3, doi: 10.1109/NSS/MIC42101.2019.9059965.

[49] L. Ma, H. Xue, G. Yang, Z. Zhang, C. Li, Y. Yao, and Y. Teng, "SCRDN: Residual dense network with self-calibrated convolutions for low dose CT image denoising," *Nucl. Instrum. Methods Phys. Res. A, Accel. Spectrom. Detect. Assoc. Equip.*, vol. 1045, Jan. 2023, Art. no. 167625, doi: 10.1016/j.nima.2022.167625.

[50] D. Kang, P. Slomka, R. Nakazato, J. Woo, D. S. Berman, C.-C.-J. Kuo, and D. Dey, "Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm," *Proc. SPIE*, vol. 8669, pp. 671–676, Mar. 2013.

[51] M. Diwakar and P. Singh, "CT image denoising using multivariate model and its method noise thresholding in non-subsampled shearlet domain," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101754, doi: 10.1016/j.bspc.2019.101754.

[52] Y. Chen, X. Yin, L. Shi, H. Shu, L. Luo, J.-L. Coatrieux, and C. Toumoulin, "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing," *Phys. Med. Biol.*, vol. 58, no. 16, p. 5803, Aug. 2014.

[53] M. Geng, X. Meng, J. Yu, L. Zhu, L. Jin, Z. Jiang, B. Qiu, H. Li, H. Kong, J. Yuan, K. Yang, H. Shan, H. Han, Z. Yang, Q. Ren, and Y. Lu, "Content-noise complementary learning for medical image denoising," *IEEE Trans. Med. Imag.*, vol. 41, no. 2, pp. 407–419, Feb. 2022, doi: 10.1109/TMI.2021.3113365.

[54] D. Wang, F. Fan, Z. Wu, R. Liu, F. Wang, and H. Yu, "CTformer: Convolution-free token2token dilated vision transformer for low-dose CT denoising," *Phys. Med. Biol.*, vol. 68, no. 6, Mar. 2023, Art. no. 065012, doi: 10.1088/1361-6560/acc000.

MUHAMMAD ZUBAIR received the Bachelor of Science (B.Sc.) degree in computer science from Peshawar University, Pakistan, the Master of Computer Science (M.C.S.) degree from Gomal University, Dera Ismail Khan, and the Master of Science (M.S.) degree from Bacha Khan University, Charsadda, Pakistan. He is currently pursuing the Ph.D. degree with Universiti Teknologi PETRONAS (UTP), Malaysia. With a strong academic foundation and a passion for innovation, his research and professional endeavors focus on advanced areas, such as machine learning, deep learning, computer vision, medical image processing, and algorithm optimization. He is also deeply engaged in feature selection and classification techniques, ensuring efficient and accurate data analysis. Additionally, he is CCIP-certified in Routing and Switching from CISO, demonstrating his expertise in the field. His interests further include computer networks, with a focus on applying intelligent algorithms to enhance network performance and connectivity.

HELMI MD RAIS received the B.Sc. degree in business administration from Drexel University, USA, the M.Sc. degree in information technology from Griffith University, Australia, and the Ph.D. degree in science and system management from Universiti Kebangsaan Malaysia (UKM). He is currently a Senior Lecturer with Universiti Teknologi PETRONAS. His research interests include ant colony algorithm, optimization, swarm intelligence, and database technologies.

TALAL ALAZEMI received the Bachelor of Engineering (B.Eng.) degree in electronics and electrical engineering from the University of Huddersfield, U.K., the Master of Science (M.Sc.) degree in power distribution engineering from Newcastle University, U.K., and the Doctor of Philosophy (Ph.D.) degree in electronic and electrical engineering from Brunel University London. His research expertise lies in renewable energy systems, artificial intelligence applications in power systems, and wind power forecasting. Recently, he has also developed a keen interest in exploring the field of image analysis, with a particular focus on medical imaging, aiming to leverage advanced computational techniques to address challenges in healthcare and diagnostics.