Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A Hybrid U-Net Framework for Low-Dose CT Image Denoising: Leveraging EdgeNet+ for Structural Integrity and Image Quality

Muhammad Zubair[1,*], Helmi Md Rais[1,2], Fasee Ullah[1], Talal Alazemi[3], Arsalaan Khan Yousafzai[4,5]

[1] Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia
[2] Institute of Health and Analytics (IHA), Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia
[3] Department of Electronic & Electrical Engineering, Brunel University London, London, Uxbridge UB8 3PH, United Kingdom
[4] Civil Engineering Department, University of Engineering of Technology Peshawar, 25120 Peshawar, Pakistan
[5] Department of Civil & Environmental Engineering, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia

ABSTRACT

Computed Tomography (CT) is a non-invasive imaging modality widely used for the precise detection of abnormalities within the human body. However, the electromagnetic radiation generated during CT scans poses health risks, including metabolic disturbances and genetic mutations, which can increase cancer susceptibility. To mitigate these hazards, Low-Dose CT (LDCT) techniques were introduced, significantly reducing radiation exposure. However, LDCT compromises image quality by introducing increased noise, artifacts, reduced contrast and structural distortions, which can impair the accuracy and reliability of Computer-Aided Diagnosis (CAD) systems. This study presents EdgeNet+, an advanced U-Net architecture with 21 convolutional layers and three skip connections, which facilitate the maintenance of high-quality structural details by allowing features from earlier layers to directly influence later stages. EdgeNet+ enhances performance by incorporating a multi-level edge detection block and a hybrid loss function. The edge detection block effectively captures edge features across various scales. The hybrid loss function combines Structural Similarity Index Measure (SSIM) and L1 losses, where SSIM promotes the preservation of image structures and perceptual quality, while L1 loss ensures accurate pixel-wise reconstruction. This combination enables EdgeNet+ to produce denoised images that are visually appealing and faithful to the original content, making it highly effective in enhancing detection accuracy. In the ablation study, the impact of each component of the EdgeNet+ model was systematically assessed, demonstrating a significant improvement in denoising performance. Comparative analysis reveals a noteworthy 45.12% increase in Peak Signal-to-Noise Ratio (PSNR), a 26.17% enhancement in SSIM and a remarkable 90.71% reduction in Root Mean Square Error (RMSE) when juxtaposed with the LDCT image. Compared to benchmark algorithms, the proposed approach demonstrates a marked improvement in noise reduction and artifact removal. Qualitative comparisons reveal a high similarity between the denoised CT images produced by EdgeNet+ and normal-dose CT images.

Keywords:
Computer tomography; medical images; multi-level edge detection; skip connections; U-Net architecture

[*] *Corresponding author.*
E-mail address: muhammad_22000228@utp.edu.my

https://doi.org/10.37934/araset.64.1.121136

1. Introduction

Medical imaging is essential for diagnosing various diseases, starting with the discovery of X-rays by Wilhelm Conrad Roentgen in 1895, which established X-ray radiography as the first diagnostic imaging modality [1]. Further, technologies such as Computed Tomography (CT), ultrasound, Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) have advanced and are now vital in-patient care, aiding in detection, characterization, staging and treatment assessment [2]. The increasing sophistication of imaging techniques has led to a significant rise in the volume of images, challenging radiologists to maintain efficient workflows [3]. The emergence of Artificial Intelligence (AI) and Deep Learning (DL) has greatly impacted medical image processing, enhancing the analysis of imaging data and improving diagnostic outcomes [4]. AI in medical imaging could potentially save 40,000 lives annually, reduce healthcare costs by over 200 billion euros and free up 1.8 billion hours of medical professionals' time, equivalent to adding 500,000 full-time healthcare workers [5].

In comparison to other medical imaging modalities, CT scans offer superior tissue differentiation and provide detailed three-dimensional images, effectively addressing the limitations of conventional X-ray techniques [6]. Additionally, CT scans can visualize both soft tissues and blood vessels simultaneously, facilitating the early diagnosis of various pathological conditions, including malignancies, vascular disorders, internal injuries, lung nodules and bone fractures. The CT scan has many contributions to medical treatment; However, ionized radiation is one of the main pitfalls, particularly for patients who undergo multiple scans [7]. According to Zubair *et al*., [5] the likelihood of developing thyroid cancer is significantly increased, with an Odds Ratio (OR) of 2.55 and a 95% Confidence Interval (CI) ranging from 2.36 to 2.75. Similarly, the risk of leukaemia is also heightened, with an OR of 1.55 and a 95% CI between 1.42 and 1.68. To address these issues, As Low as Reasonably Achievable (ALARA) principle has been adopted and is a fundamental guideline in the CT imaging community to minimize radiation exposure to patients [8,9]. To mitigate the radiation exposure associated with CT scans, LDCT technology has been developed. This approach utilizes two primary strategies: first, by decreasing the flux of the X-ray tube and second, by reducing the number of scan trajectories. Both methods ultimately lead to a diminished Signal-to-Noise Ratio (SNR) for the X-ray signals [10,11]. However, this introduced noise and artifacts in CT images, which can lower the SNR and impact the diagnostic accuracy [12]. Removing these noise and artifacts from LDCT images poses a significant challenge.

Numerous algorithms have been developed to enhance the quality of LDCT images. These algorithms are generally classified into three groups: Sinogram filtering, Iterative Reconstruction (IR) and Post-processing techniques [13,14]. Where Sinogram filtering, preprocessing the raw data before CT image reconstruction, utilizes methods like Filtered Back-Projection (FBP). One advantage of this approach is its reliance on well-known noise characteristics in the Sinogram domain [15]. However, Sinogram-based filtering methods may suffer from drawbacks such as spatial resolution loss or edge blurring, thereby limiting their effectiveness [16]. IR methods are continuously improved by refining objective functions, considering the statistical aspects of projection data, real CT image insights and system-specific parameters [17]. These methods encompass dictionary learning, nonlocal means, low-rank, total variance and other approaches [18,19]. Many modern Multidetector Computed Tomography (MDCT) scanners have adopted IR techniques, which have demonstrated significant advancements. However, they possess two primary limitations: First, the proprietary nature of IR techniques limits their interoperability across different scanner manufacturers due to restricted access to scanner geometry and correction procedures. Secondly, widely used IR techniques incur substantial computational overheads [20]. The third category of LDCT image denoising is post-

processing methods. These methods operate directly on the CT images, requiring no prior knowledge of raw data and offering practical convenience [21,22]. Post-processing methods are generally divided into traditional techniques and deep learning-based approaches. Traditional methods, relying on predefined noise-image relationships, employ optimization algorithms and prior noise knowledge for denoising. Further, wavelet-based denoising filters denoise the CT image but may blur fine details, while Block-Matching and 3D Filtering (BM3D) filtering enhances anatomical visualization yet struggle with streak artifacts near bones [23]. Dictionary learning effectively removes Gaussian noise but has limited success with complex LDCT noise and Non-Local Means (NLM) algorithms preserve features well but are computationally intensive. Although traditional methods can be costly and may introduce artifacts in uniform areas, deep learning-based methods offer more efficient and effective solutions to improve LDCT image quality [6]. Deep learning techniques have shown considerable success in denoising CT images, effectively reducing noise and improving image quality. Further, deep learning-based image-denoising models are divided into three categories Discriminative models (CNN and Its variants), Generative models (Autoencoder and U-Net Models) and Hybrid Models (Combination of Discriminative and Generative models). Unlike conventional techniques reliant on prior information, deep learning, notably Convolutional Neural Networks (CNN), autonomously learns intricate image representations, excelling in extracting features from CT images and accommodating diverse noise patterns [24]. Further, Zubair *et al*., [10] introduced a deep learning model, DEPnet, utilizing the dilated convolution with a batch normalization layer. Moreover, Shan *et al*., [25] introduced MAP-NN, which employs progressive denoising through multiple CLONEs and optimization with composite loss functions, however, this approach may lead to heightened computational demands and longer training durations. Further, Gou *et al.,* [26] proposed GRCNN, which integrates both pixel-wise grey value loss and image gradient loss, providing a comprehensive approach to preserve image details during denoising. Furthermore, Ansari *et al.,* [27] combined an edge detection layer and perceptual loss to denoise the LDCT image. Additionally, Liang *et al.,* [28] proposed EDCNN for LDCT image denoising challenges using trainable Sobel convolution for edge enhancement and dense connections for feature fusion, however, it needs computational resources during training and inference. Moreover, Liu *et al.,* [29] proposed Stacked Sparse Denoising Autoencoders (SSDAs) to enhance the quality of LDCT imaging, however, the proposed methodology has a potential risk of overfitting. Also, Fenga *et al.,* [30] introduced a novel DRCNN model which combines MSE losses for sinogram and image domains, along with TV regularization. Additionally Feng *et al.,* [31] combined residual U-Net with a multidimensional feature extraction unit and spatial attention mechanisms, however, it faces challenges in generalizing its performance across different datasets and imaging conditions. Further, Zhang *et al.,* [32] combined U-Net architecture with a multi-attention mechanism, comprising three attention modules for contextual information extraction, feature extraction and detail retention, but computationally complex. Furthermore, Jiao *et al.,* [33] used an encoder-dual decoder sub-network to reduce noise but may inadvertently create artifacts or distortions in the processed images, which can degrade the overall image quality. Further, Zubair *et al.,* [21] introduced a DoG-UNet+ deep learning algorithm that enhanced U-Net with a "Difference of Gaussians Sharpening Layer" and utilized dual convolutional kernels to capture diverse features. Further, this technique used an attention mechanism to highlight important features in CT images, its performance warrants validation on real datasets. Also, Park *et al.,* [34] introduced a fidelity-embedded GAN to learn from unpaired LDCT and SDCT images, optimizing a generator by minimizing a weighted sum of KL divergence and loss. Further, Chi *et al.,* [35] proposed a methodology using an enhanced U-Net and a novel discriminator with multiple CNNs to denoise LDCT images. One area that has been comparatively underexplored in the literature is features edge detection, a critical aspect of medical imaging. Edge detection significantly enhances the ability to

identify anatomical structures with high precision, directly contributing to improved diagnostic accuracy and more effective treatment planning. This paper has the below-mentioned contributions.

i. EdgeNet+ introduces a refined U-Net architecture tailored for CT image denoising utilizing 21 convolutional layers and three skip connections to enhance feature preservation and denoising effectiveness.
ii. A novel edge detection block is introduced, combining convolutional layers with multiple kernel sizes (3x3, 5x5, 7x7 and 9x9) and a residual connection to capture edges across various scales.
iii. A hybrid loss function is proposed, combining SSIM loss to preserve structural information and L1 loss to ensure smooth gradient signals. This combination aims to maintain both perceptual quality and structural integrity in denoised images.
iv. An ablation study is undertaken on the EdgeNet+ model to rigorously assess, the significance and functionality of its components and mechanisms.
v. The EdgeNet+ model employs a hybrid loss function and is evaluated on the COVID-19 artificial LDCT dataset using performance metrics, including PSNR, SSIM and RMSE.
vi. The performance of the EdgeNet+ algorithm is evaluated against established state-of-the-art models by analysing regions of interest (ROIs) from two images in the COVID-19 LDCT test dataset.

This paper is structured as follows: Section 2 presents the “Methodology”, providing a detailed description of the approach employed in this study. Section 3, titled "Experiments and Results", evaluates the proposed model through empirical analysis. Finally, Section 4, "Conclusion and Future Work", summarizes the key findings and proposes potential avenues for future research.

2. Methodology

2.1 Objective Function

2.1.1 Structure similarity loss

Enhancing the method of minimizing the least amount of square loss can result in denoised images that closely resemble the distribution found in normal-dose CT scans. However, it often fails to retain intricate details. The typical Mean Squared Error (MSE) loss employed in CNN-based techniques often results in excessively smooth and blurred images, causing a degradation of structural details [36]. In medical imaging, CT images exhibit variations in dose levels. The relationships between features in CT images are significant. The SSIM considered three key aspects of human visual perception: luminance, contrast and structure. SSIM demonstrates superior performance compared to conventional loss functions such as MSE in diagnostic accuracy [37]. To quantify the similarity between a denoised CT image and its normal-dose counterpart, we are employing a hybrid loss function. One component of this loss function is to formulate and assess the resemblance between CT images, based on their luminance, contrast and structural similarities as given in Eq. (1).

$$SSIM\ (x, y) = \frac{2\mu_x\mu_{y+C_1}}{\mu_x^2+\mu_y^2+C_1} * \frac{\sigma_{xz+C_2}}{\sigma_x^2+\alpha_y^2+C_2} \quad (1)$$

where the parameters *"μ_x", "μ_y", "σ_x", "σ_y" and "σ_{xy}"* denote the means, standard deviations and cross-correlation of two CT images being compared. *"C1"* and *"C2"* are constants utilized to prevent

numerical issues. As the similarity between x and y grows, the SSIM value tends towards one. Hence, the loss function for SSIM can be expressed as Eq. (2).

$Loss_{SSIM}$*=1- SSIM (x, y)* (2)

2.1.2 Least absolute error

The "L1" loss, also known as the Least Absolute Error (LAE), serves as a loss to evaluate the difference between two images. Unlike the MSE loss, the "L1" loss does not overly penalize larger discrepancies between a denoised image and its ground truth counterpart. This indicates that in scenarios where substantial discrepancies exist, the "L1" loss treats them more leniently compared to the MSE loss. One of the notable advantages of the "L1" loss is its ability to mitigate certain drawbacks often associated with the MSE loss [38]. These drawbacks include the tendency for MSE to produce blurred outputs and introduce unnatural artifacts. By utilizing the "L1" loss, these issues can be alleviated to some extent as given in Eq. (3).

$L1_{LOSS} = \frac{1}{mnb}| x-y|$ (3)

where, *"x"* and *"y"* represent the noisy and original CT images, while *"m"* and *"n"* stand for the dimensions of the CT image and *"b"* indicates the batch size.

2.1.3 Hybrid loss function

The comprehensive objective function incorporates the least squares, *"L1" and* SSIM to ensure denoised CT images of high quality with preserved texture and structure as given in Eq. (4).

$HYBRIDLOSS= \alpha(1 - \frac{2\mu_x\mu_{y+C_1}}{\mu_x^2+ \mu_y^2+ C_1} * \frac{\sigma_{xz+C_2}}{\sigma_x^2+ \alpha_y^2+ C_2})+ \beta(\frac{1}{mnb} *| x-y|)$ (4)

where *α* and *β* represent the weights assigned to prioritize different components.

2.2 EdgeNet+ Architecture

The EdgeNet+ architecture is characterized by its symmetric structure, which contains an encoder, latent space, Multi-Level Edge Detection Block and a decoder.

2.2.1 Encoder

Let h (0) = X, which denotes the input grayscale CT images with dimensions 256x256. The convolutional hidden layers are represented as mentioned in Eq. (5).

$h(l) = Activation(W(l) * h(l-1) +b(l))$ (5)

where "Activation" represents the activation function, *"W"* and *"b"* signify the weights and biases for the layers respectively. Also, "h" represents hidden layers and "l" denotes the number of layers, ranging from 1 to "l-1". These convolutional layers have "3×3" filters to extract features from the input CT images. After each pair of convolutional layers, max-pooling layers with a "2×2" window size

are applied to down-sample the feature maps, reducing their spatial dimensions while preserving important features as given in Eq. (6).

h(c)= (h(l), k) (6)

where *"K"* indicates the Max pooling layer as applied on the feature map generated by the convolutional layer *"(h(l))".*

2.2.2 Bottleneck layers

The bottleneck layer serves as a bottleneck for feature extraction, where the most abstract and high-level features are encoded. If *"Z"* denotes the output feature maps of the latent space. Mathematically, the bottleneck layer can be represented as given in Eq. (7).

Z= (h(c)) (7)

where *"h(c)"* indicates the output feature maps of the last convolutional layer in the encoder.

2.2.3 Multi-level edge detection block

The EdgeNet+ model strategically incorporates the Multi-Level Edge Detection Block at the start of the decoder. This addition is pivotal for capturing edge features across multiple scales during feature extraction. This novel Multi-Level Edge Detection Block utilizes four convolutional layers with varying kernel sizes: "3x3", "5x5", "7x7" and "9x9" to process the CT image. Each convolutional layer is followed by a Rectified Linear Unit (ReLU) activation function to introduce non-linearity and a skip connection links the input of the block to this layer's output. By setting the padding parameter to 'same', the output feature maps maintain the same spatial dimensions as the input. The output feature maps from the convolutional layers are then concatenated along the channel axis using the concatenate operation. This concatenation combines edge information captured at different levels, enhancing the model's edge detection capabilities. The output of the latent space is "Z", then mathematically it can be represented in Eq. (8).

$$C(i,j)= Activation\left(\sum_{u=0}^{P^{(i)},-1)} \sum_{v=0}^{q^{(i)}-1)} Z^{(l-1)}(i+u, J+\ v) * \ K^{l}(u,v) + b^{l}\right) \quad (8)$$

where, *"C(i,j) "* is the output feature map at position *"(i,j) "* in the *"l-th"* convolutional layer. *"Z* "signifies the feature map from the latent space. The K indicates the Kernal matrix for the *"l-th"* convolutional layer. Also, *"$b^{(l)}$ "*represents the bias for the *"l-th"* convolutional layer. The Kernal size "*$p^{(i)}$ $q^{(i)}$"* varies at each layer. The block of the four convolutional layers is backed by a Max-pooling layer as mentioned in Eq. (9).

$$P'_{(i',j')} = (C(i,j),k') \quad (9)$$

where *"k'"* is the max pooling layer, applied on the Multi-Level Edge Detection Block *"C(i,j)"*, as illustrated in Figure 1.

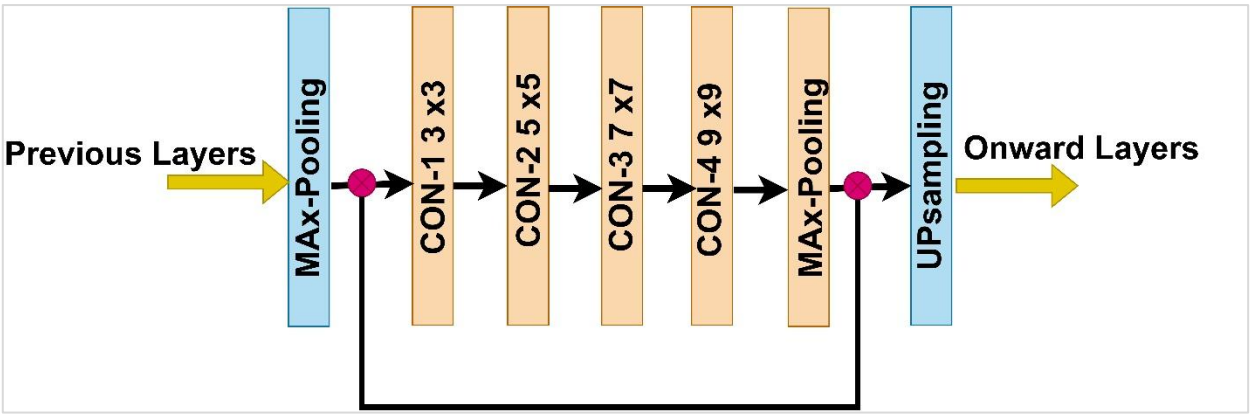


Fig. 1. Multi-level edge detection block

2.2.4 Decoder

The output of the Multi-Level Edge Detection Block serves as the input to the Up-sampling layer in the decoder section. Where the up-sampling layer expands the spatial dimensions of the feature maps to align with those of the corresponding feature maps in the encoder section, while the convolutional layers refine the features obtained from the Up-sampling layers, capturing finer details and patterns. The ReLU activation functions introduce non-linearity to extract abstract representations. The mathematical representation of the decoder part is given in Eq. (10).

$$h^{(l)} = Activation(w'^{l} * h^{(l-1)} + b') \quad (10)$$

2.2.5 Skip connections

The proposed EdgeNet+ model contains three concatenation layers. The convolutional layers are concatenated with their corresponding Up-sampling layers to facilitate the reconstruction process, where the Up-sampling Layer 3 relates to Convolutional Layer 2 and Up-sampling Layer 2 relates to Convolutional Layer 4. Finally, Convolutional Layer 6 is concatenated with Up-sampling Layer 1, enabling the fusion of the most abstract features from the convolutional layer with the Up-sampled feature maps at the highest spatial resolution, facilitating the generation of the final reconstructed output with the desired level of detail and fidelity [39]. These concatenation operations merge feature maps with the same spatial dimensions along the channel axis, allowing the model to combine low-level and high-level features effectively. This integration enables the EdgeNet+ model to capture local and global contextual information, facilitating accurate denoising of LDCT images while preserving fine details. Mathematically the skip connection can be represented as given in Eq. (11).

$$Y(i) = f(x)i + xi \quad (11)$$

where I =1,2,3,4. The proposed EdgeNet+ diagram is given in Figure 2.

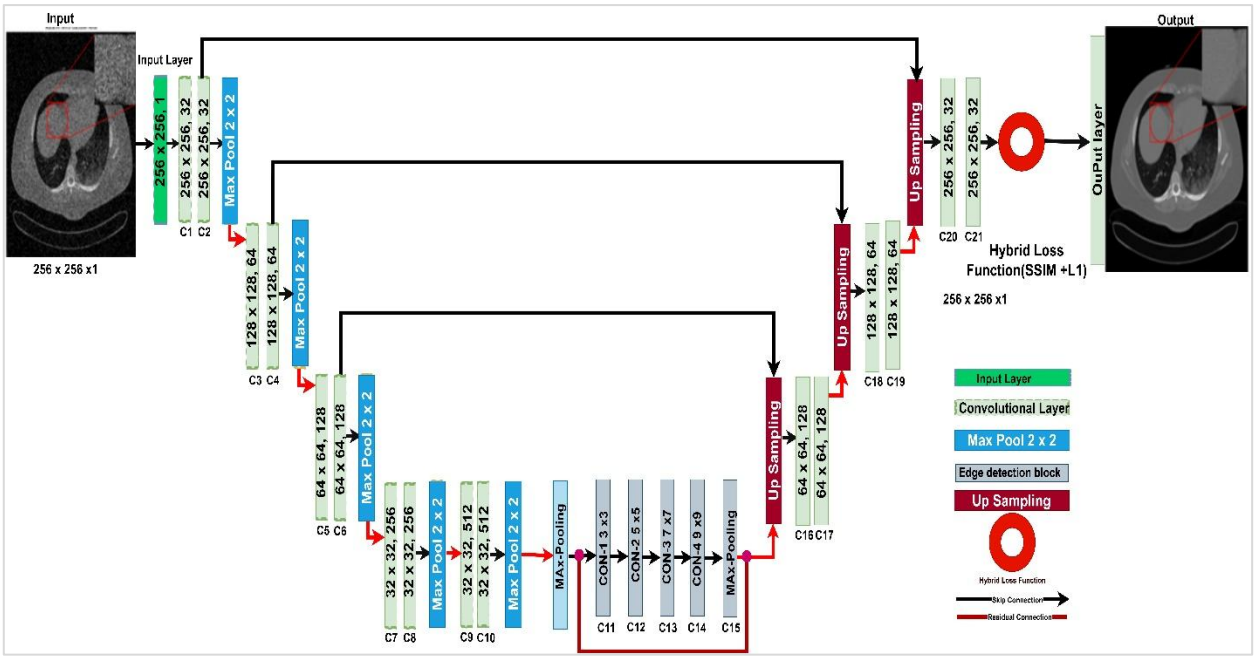


Fig. 2. Proposed EdgeNet+ diagram

2.3 Model Evaluation

The proposed model, EdgeNet+, has assessed state-of-the-art methods using metrics such as PSNR, SSIM and RMSE to evaluate performance.

2.3.1 PSNR

PSNR is frequently employed as a quantitative measure to assess the denoising efficacy of CT images. It quantifies the quality of the denoised image by comparing it to the original image, considering both the signal strength and the presence of noise. Higher PSNR values indicate better denoising performance, with increased fidelity in preserving image details and reduced distortion as given in Eq. (12).

$$PSNR = 10.log_{10}(\frac{MAX^2}{MSE}) \tag{12}$$

Where, "MAX" is the maximum possible pixel value of the image (typically 255 for 8-bit images) and MSE is the Mean Squared Error between the original and denoised images, calculated as the average of the squared differences between corresponding pixels.

2.3.2 RMSE

RMSE is a precise metric used to measure the average difference between pixel intensities in the denoised image and the corresponding ground truth LDCT image. It is calculated as given in Eq. (13).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{denoised}\,(i) - I_{LDCT}(i)^2} \tag{13}$$

where *"N"* is the total number of pixels in the image. "$I_{denoised}(i)$"and "$I_{LDCT}(i)$" represent the intensity values of the denoised image and the ground truth LDCT image, respectively, at pixel i.

2.3.3 SSIM

SSIM metric provides a comprehensive assessment of how well the denoised image preserves important structural information compared to the ground truth CT image [40]. A higher SSIM value implies that the denoised image closely mirrors the ground truth CT image, demonstrating superior denoising performance. SSIM can represented as given in Eq. (14).

$$SSIM\ (X1, X2) = \frac{(2\mu_{x_1}\mu_{x_2} + C_1)}{(\mu_{x_1}^2 + \mu_{x_2}^2 + C_1)} \frac{(2\alpha_{x_1x_2} + C_2)}{(\alpha_{x_2}^2 + \alpha_{x2}^2 + C_2)} \quad (14)$$

where *"x"* and *"y"* are the input images: ground truth and denoised images to compare and "μ_x","$a_{x,}$" and "$\alpha_{x_1\,x_2}$ " which signify the means, standard deviation and cross-correlation of two patches, respectively [41]. Additionally, "C_1" and "C_2" serve as positive constants to stabilize calculations in the context of the SSIM formula.

3. Experiments and Results

3.1 Dataset for Model Evaluation

To evaluate the effectiveness of the proposed EdgeNet+ model, a simulated COVID-19 dataset was utilized available publicly on GitHub. The dataset is partitioned into three subsets. The training set comprises 2097 pairs of images, each measuring 256x256 pixels, constituting 80% of the total dataset. Each pair consists of a clean image and its corresponding noisy version. The validation and test sets consist of 260 noisy images each, mirroring the dimensions of the training set. They each account for 10% of the dataset. This dataset has been meticulously organized to facilitate the training, validation and assessment phases of the proposed EdgeNet+ model. During the training phase, the model is trained using the training set. The validation set is utilized to fine-tune hyperparameters and evaluate the model's performance during training. Finally, the test set remains reserved for the final evaluation, assessing the model's ability to generalize to unseen data.

3.2 Parameters Setting

The model is trained using the Adam optimizer with a learning rate of 1e-3. The coefficients of the objective function (α=0.005, β=0.0093) are empirically determined values. These coefficients are crucial in the custom loss function, which combines SSIM loss and L1 loss. Adjusting α and β allows for control over balancing the measurement of structural similarity and absolute difference in the output. The performance of the model is evaluated using PSNR, RMSE and SSIM metrics. The input images are resized to 256x256 pixels and the training is conducted using a batch size of 32 for 100 epochs. Additionally, the model architecture includes a multi-level edge detection block comprising convolutional layers with kernel sizes of "3x3", "5x5", "7x7" and "9x9", followed by a max-pooling layer. The experiments were conducted using a personal computer equipped with an Intel i5-4300 processor and 16 GB of RAM. The computational power was boosted by NVIDIA Tesla T4 GPU acceleration.

3.3 Qualitative Results

To assess the efficacy of the proposed EdgeNet+ approach, two distinct LDCT images have been selected from the COVID-19 test dataset. Within these images, two Regions of Interest (ROIs) have been identified for comparison purposes. The performance of the proposed method, EdgeNet+ has been compared, with four state-of-the-art algorithms: NLM, BM3D, RED-CNN and DOG-UNet+. The denoised results obtained from these methods are illustrated in Figure 3 and Figure 4, Where NLM and BM3D are two popular traditional denoising schemes extensively utilized for denoising CT images, While the RED-CNN utilizes an autoencoder approach with MSE cost function. On the other hand, the DOG_UNet+ model adopts a deep learning autoencoder methodology, incorporating a custom loss function alongside an attention mechanism.

3.3.1 LDCT first image (ROI-I)

Figure 3 presents CT images alongside their zoomed-in versions to assess the performance of various algorithms. The first CT image from the COVID-19 test dataset is depicted.

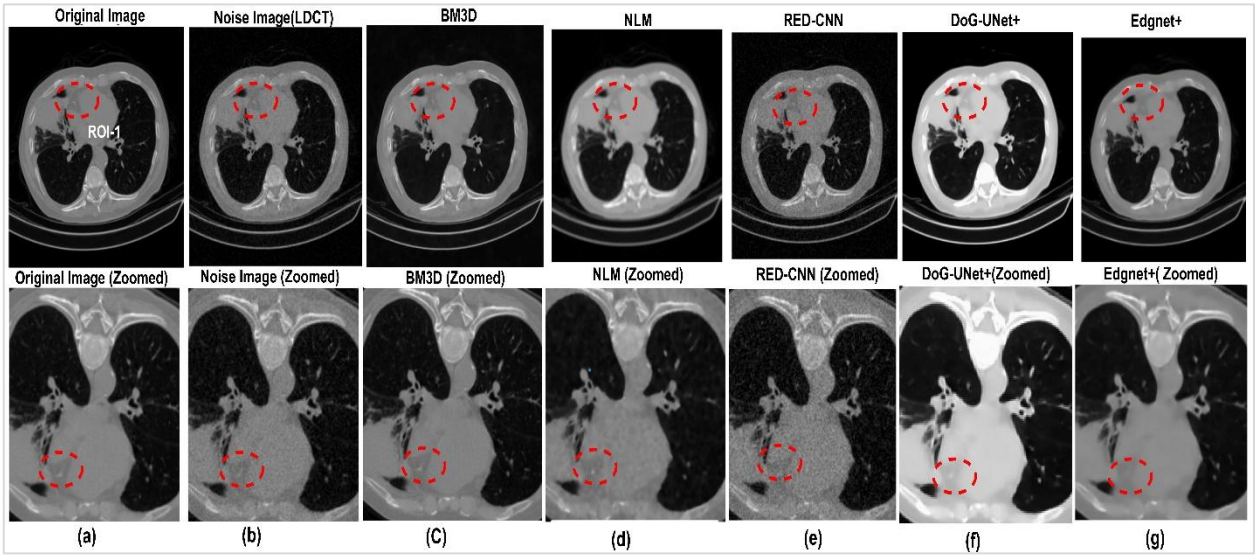


Fig. 3. Results from different methods, region of interest (ROI-I) along with zoomed images version (a) original image (b) LDCT image (c) BM3D (d) NLM (e) RED-CNN (f) DOG-UNet

Where Figure 3(a) illustrates the original CT image and Figure 3(b) presents the corresponding LDCT image. To assess the performance of different denoising techniques, a specific Region of Interest (ROI-I) is carefully chosen. When traditional algorithms like BM3D and NLM are applied to the LDCT image Figure 3(b), BM3D effectively preserves edges and clarity, as shown in Figure 3(c), while NLM results in blurring and over-smoothing effects, making it difficult to read the ROI as depicted in Figure 3(d). Consequently, BM3D in Figure 3(c) exhibits superior performance over NLM in Figure 3(d) in terms of enhancing visual quality. Transitioning to deep learning methods, including RED-CNN, DOG_UNet+ and the proposed EdgeNet+, it is observed that RED_CNN Figure 3(e) denoises the LDCT image Figure 3(b) while maintaining edge information, but it introduces streak artifacts. Conversely, DOG_UNet+ Figure 3(f) preserves edge information but exhibits an over-smoothing effect, showing superior performance compared to RED_CNN Figure 3(e). Finally, when the proposed method, EdgeNet+, is applied to LDCT images, the denoised results in Figure 3(g) not only preserve edge information but also exhibit superior visual quality compared to all other methods.

3.3.2 LDCT second image (ROI-II)

The second CT image from the COVID-19 test dataset is selected. Image Figure 4(a) displays the original CT image, while Figure 4(b) shows the corresponding LDCT image. To evaluate the efficacy of our proposed method compared to other denoising methods, precise ROI-II is selected.

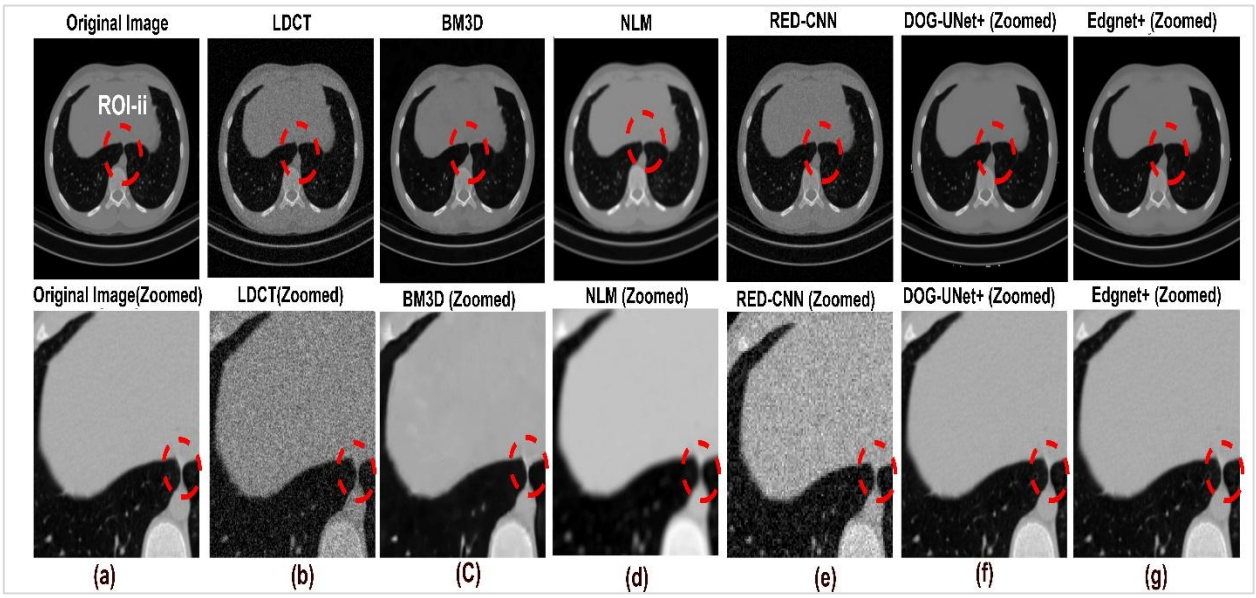


Fig. 4. Results from different methods, region of interest (ROI-II) along with zoomed images version (a) original image (b) LDCT image (C) BM3D (d) NLM (e) RED-CNN (f) DOG-UNet

Where BM3D and NLM are both widely recognized traditional methods for denoising CT images. When utilized on the LDCT image in Figure 4(b), both methods yield enhanced outcomes, however, in Figure 4(c), BM3D preserves the clarity, while NLM in Figure 4(d) introduces some blurriness into the denoised output. After analysing the experimental findings, it was observed that RED-CNN Figure 4(e) effectively preserves the clarity of CT images, albeit with the drawback of introducing blocky artifacts. Conversely, DOG_Unet+ demonstrates superior performance compared to RED-CNN, showcasing better results. Finally, when the proposed EdgeNet+ methods were applied to LDCT images, the results depicted in Figure 4(g) for EdgeNet showcased not only improved denoising but also remarkable preservation of image clarity.

3.4 Quantitative Results

Three widely used metrics PSNR, SSIM and RMSE were employed for quantitative assessments of Regions of Interest (ROI-I and ROI-II). PSNR and RMSE evaluate denoising efficacy at the pixel level, while SSIM measures structural similarity within a defined window. The quantitative results from various denoising methods are presented in Table 1 and Table 2, corresponding to the LDCT images depicted in Figure 3 and Figure 4. The first two methods, Block-Matching and 3D Filtering (BM3D) and Non-Local Means (NLM), represent traditional approaches. In contrast, RED-CNN, DOG_UNet and EdgeNet+ are deep learning methods utilizing distinct cost functions for CT image denoising. Notably, EdgeNet demonstrated superior performance compared to the other methods.

Table 1
Quantitative results of ROI-I from Figure -3 in terms of PSNR, SSIM and RMSE metrics

S/No	Protocols	PSNR	SSIM	RMSE
1	LDCT	27.125	0.818	0.136
2	BM3D	32.913	0.899	0.028
3	NLM	31.241	0.898	0.029
4	RED-CNN	28.435	0.813	0.049
5	DOG-UNet+	35.355	0.878	0.026
6	EdgeNet+ (Proposed)	38.763	0.973	0.014

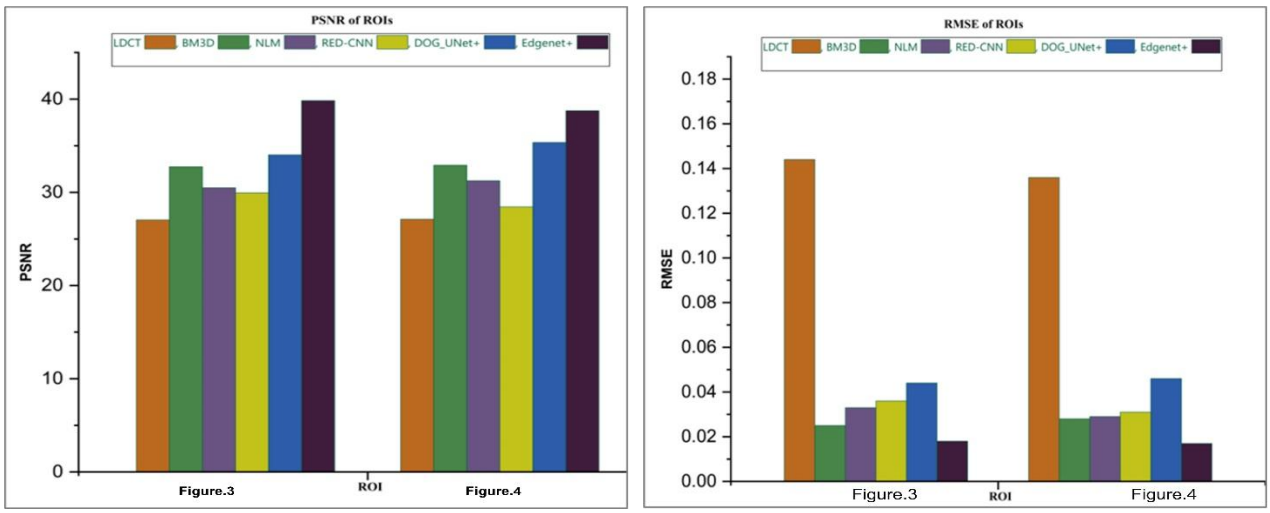
Table 2
Quantitative results of ROI-II from Figure-4 in Terms of PSNR, SSIM and RMSE

S/No	Protocol	PSNR	SSIM	RMSE
1	LDCT	27.041	0.725	0.144
2	BM3D	32.744	0.821	0.025
3	NLM	30.495	0.934	0.033
4	RED-CNN	29.942	0.769	0.046
5	DOG-UNet+	34.022	0.824	0.021
6	EdgeNet+ (Proposed)	39.841	0.974	0.013

Table 3
Average quantitative results from Table 1 and Table 2 in Terms of PSNR, SSIM and RMSE

S/No	Protocol	PSNR	SSIM	RMSE
1	LDCT	27.083	0.772	0.140
2	BM3D	32.829	0.860	0.026
3	NLM	30.868	0.916	0.031
4	RED-CNN	29.189	0.791	0.047
5	DOG-UNet+	34.689	0.851	0.023
6	EdgeNet+ (Proposed)	39.302	0.974	0.013

Figure 5 reflects the ROIs in Figure 3 and Figure 4.



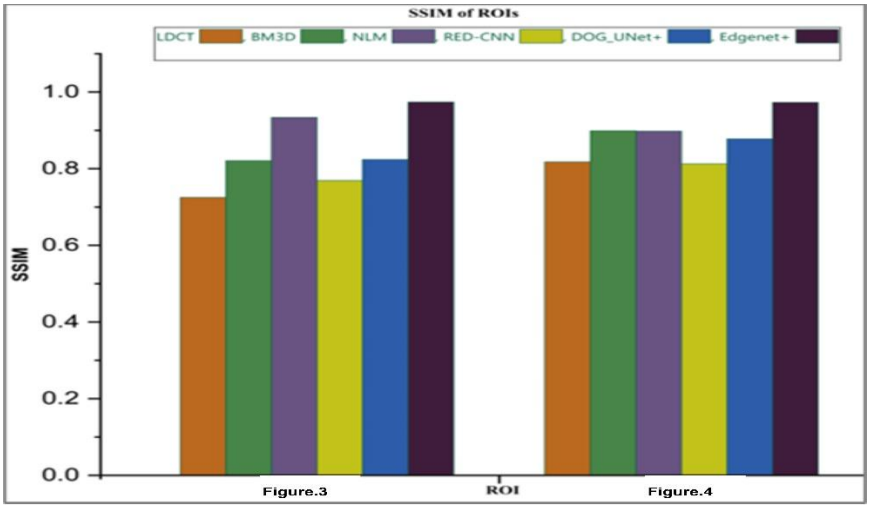


Fig. 5. ROIs comparison of Figure 3 and Figure 4

4. Ablation Study

In this section, an ablation study was conducted to comprehensively analyse the effectiveness of the proposed method EdgeNet+. In the component analysis, we delve into the impacts of three configurations: the basic U-Net model, the U-Net model coupled with the $HYBRID_{LOSS}$ function and the U-Net model augmented with both $HYBRID_{LOSS}$ and the Multi Edge Detection Block as mentioned in Table 4.

Table 4
Ablation study of the component analysis of EdgeNet+ model

Method	PSNR	SSIM	RMSE
U.Net model	36.231	0.967	0.252
U.Net model + HYBRIDLOSS	37.472	0.971	0.181
U.Net model+ HYBRIDLOSS + Multi Edge Detection Block	39.841	0.972	0.141

To evaluate the impact of different components of the proposed model EdgeNet a notable improvement across several metrics was observed. Initially, we employed the basic U-Net model along with four skip connections that yielded a PSNR of 36.231, SSIM of 0.967 and RMSE of 0.252. Upon integrating the $HYBRID_{LOSS}$ function, these metrics demonstrate enhancements, with PSNR increasing to 37.472, SSIM to 0.971 and RMSE decreasing to 0.181. Further augmentation of the Multi Edge Detection Block yields substantial improvements, resulting in a PSNR of 39.841, SSIM of 0.972 and RMSE of 0.141. This sequential enhancement underscores the synergistic contribution of both the HYBRIDLOSS and the Multi Edge Detection Block in refining the denoising capabilities of the EdgeNet+ model, ultimately yielding significant advancements in PSNR, SSIM and reduction in RMSE.

5. Conclusion

In conclusion, the EdgeNet+ model demonstrates significant advancements in denoising low-dose CT images by effectively integrating skip connections, a Multi Edge Detection Block and a hybrid loss function. The model's ability to preserve intricate details and edge clarity sets it apart from traditional and contemporary state-of-the-art algorithms. Comparative analyses reveal that EdgeNet+ not only surpasses the performance of established methods like BM3D and NLM but also outperforms deep learning approaches such as RED-CNN and DOG_Unet+. The evaluation metrics, including PSNR, SSIM

and RMSE, underscore the superior denoising quality achieved by EdgeNet+, establishing it as a leading solution for enhancing image clarity and accuracy in medical imaging applications. The findings from the ablation study further validate the effectiveness of the proposed enhancements, indicating promising implications for future research and practical applications in image restoration. Overall, the findings support the conclusion that EdgeNet+ demonstrates significant enhancements, with a 45.12% increase in PSNR, a 26.17% improvement in SSIM and a remarkable 90.71% reduction in RMSE when compared to the LDCT image. In future studies, it will be crucial to validate the proposed method using real datasets and to explore the incorporation of attention mechanisms.

Acknowledgement
This study received funding from the Department of Computer and Information Sciences and the Institute of Health and Analytics at Universiti Teknologi PETRONAS, Malaysia. We extend our sincere gratitude to these departments for their valuable support in furthering scientific understanding and fostering research excellence.

References

[1] Rusanov, Branimir, Ghulam Mubashar Hassan, Mark Reynolds, Mahsheed Sabet, Jake Kendrick, Pejman Rowshanfarzad and Martin Ebert. "Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: A systematic review." *Medical Physics* 49, no. 9 (2022): 6019-6054. https://doi.org/10.1002/mp.15840

[2] Heshmatzadeh Behzadi, Ashkan, Zerwa Farooq, Jeffery H. Newhouse and Martin R. Prince. "MRI and CT contrast media extravasation." *Medicine* 97, no. 9 (2018): e0055-e0055. https://doi.org/10.1097/MD.0000000000010055

[3] Zhang, Minghan, Sai Gu and Yuhui Shi. "The use of deep learning methods in low-dose computed tomography image reconstruction: a systematic review." *Complex & intelligent systems* 8, no. 6 (2022): 5545-5561. https://doi.org/10.1007/s40747-022-00724-7

[4] Shao, Yu-Hsuan, Kevin Tsai, Sinae Kim, Yu-Jen Wu and Kitaw Demissie. "Exposure to tomographic scans and cancer risks." *JNCI cancer spectrum* 4, no. 1 (2020): pkz072. https://doi.org/10.1093/jncics/pkz072

[5] Zubair, Muhammad, Helmi Md Rais, Qasem Al-Tashi, Fasee Ullah, Muhammad Faheeem and Arfat Ahmad Khan. "Enabling Predication of the Deep Learning Algorithms for Low-Dose CT Scan Image Denoising Models: A Systematic Literature Review." *IEEE Access* (2024). https://doi.org/10.1109/ACCESS.2024.3407774

[6] Dodia, Shubham, B. Annappa and Padukudru A. Mahesh. "Recent advancements in deep learning based lung cancer detection: A systematic review." *Engineering Applications of Artificial Intelligence* 116 (2022): 105490. https://doi.org/10.1016/j.engappai.2022.105490

[7] Diwakar, Manoj and Manoj Kumar. "A review on CT image noise and its denoising." *Biomedical Signal Processing and Control* 42 (2018): 73-88. https://doi.org/10.1016/j.bspc.2018.01.010

[8] Yoo, Seung-Jin, Young Sik Park, Hyewon Choi, Da Som Kim, Jin Mo Goo and Soon Ho Yoon. "Prospective evaluation of deep learning image reconstruction for Lung-RADS and automatic nodule volumetry on ultralow-dose chest CT." *Plos one* 19, no. 2 (2024): e0297390. https://doi.org/10.1371/journal.pone.0297390

[9] Tzimas, Georgios, David T. Ryan, David J. Murphy, Jonathon A. Leipsic and Jonathan D. Dodd. "Cardiovascular CT, MRI and PET/CT in 2021: review of key articles." *Radiology* 305, no. 3 (2022): 538-554. https://doi.org/10.1148/radiol.221181

[10] Zubair, Muhammad, B. Helmi and Qasem Al-Tashi. "U-net autoencoder for edge-preserved denoising of low dose computed tomography images: a novel technique." In *2023 13th International Conference on Information Technology in Asia (CITA)*, pp. 19-24. IEEE, 2023. https://doi.org/10.1109/CITA58204.2023.10262803

[11] Kaur, Prabhpreet, Gurvinder Singh and Parminder Kaur. "A review of denoising medical images using machine learning approaches." *Current medical imaging* 14, no. 5 (2018): 675-685. https://doi.org/10.2174/1573405613666170428154156

[12] Bousse, Alexandre, Venkata Sai Sundar Kandarpa, Kuangyu Shi, Kuang Gong, Jae Sung Lee, Chi Liu and Dimitris Visvikis. "A review on low-dose emission tomography post-reconstruction denoising with neural network approaches." *IEEE Transactions on Radiation and Plasma Medical Sciences* (2024). https://doi.org/10.1109/TRPMS.2023.3349194

[13] May, Thet Tun and Sugiura Yosuke. "Joint Training of Noisy Image Patch and Impulse Response of Low-Pass Filter in CNN for Image Denoising." *信号処理* 28, no. 1 (2024): 1-17. https://doi.org/10.2299/jsp.28.1

[14] Khaw, Li Wen and Shahrum Shah Abdullah. "Mri Brain Image Classification Using Convolutional Neural Networks and Transfer Learning." *Journal of Advanced Research in Computing and Applications* 31, no. 1 (2023): 20-26. https://doi.org/10.37934/arca.31.1.2026

[15] Zhao, Lijun, Muhammad Shahzad Nazir, Hafiz M. Jamsheed Nazir and Ahmed N. Abdalla. "A review on proliferation of artificial intelligence in wind energy forecasting and instrumentation management." *Environmental Science and Pollution Research* 29, no. 29 (2022): 43690-43709. https://doi.org/10.1007/s11356-022-19902-8

[16] Kulathilake, KA Saneera Hemantha, Nor Aniza Abdullah, Aznul Qalid Md Sabri and Khin Wee Lai. "A review on deep learning approaches for low-dose computed tomography restoration." *Complex & Intelligent Systems* 9, no. 3 (2023): 2713-2745. https://doi.org/10.1007/s40747-021-00405-x

[17] Arndt, Clemens, Felix Güttler andreas Heinrich, Florian Bürckenmeyer, Ioannis Diamantis and Ulf Teichgräber. "Deep learning CT image reconstruction in clinical practice." In *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, vol. 193, no. 03, pp. 252-261. Georg Thieme Verlag KG, 2021. https://doi.org/10.1055/a-1248-2556

[18] Arshaghi, Ali, Mohsen Ashourian and Leila Ghabeli. "Denoising medical images using machine learning, deep learning approaches: a survey." *Current Medical Imaging* 17, no. 5 (2021): 578-594. https://doi.org/10.2174/1573405616666201118122908

[19] Suleiman, Ahmad Abubakar, Arsalaan Khan Yousafzai and Muhammad Zubair. "Comparative Analysis of Machine Learning and Deep Learning Models for Groundwater Potability Classification." *Engineering Proceedings* 56, no. 1 (2023): 249. https://doi.org/10.3390/ASEC2023-15506

[20] Mück, Jonas, Elisa Reiter, Wilfried Klingert, Elisa Bertolani, Martin Schenk, Konstantin Nikolaou, Saif Afat and Andreas S. Brendlin. "Towards safer imaging: A comparative study of deep learning-based denoising and iterative reconstruction in intraindividual low-dose CT scans using an in-vivo large animal model." *European Journal of Radiology* 171 (2024): 111267. https://doi.org/10.1016/j.ejrad.2023.111267

[21] Zubair, Muhammad, B. Helmi, Fasee Ullah, Arsalaan Khan Yousafzai and Farrukh Hassan. "Enhancing Low-Dose CT Image Quality Through Deep Learning: A DoG-Sharpened U-Net Approach With Attention Mechanism." In *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, pp. 1037-1041. IEEE, 2024. https://doi.org/10.1109/ICETSIS61505.2024.10459568

[22] Mustafa, Wan Azani, Haniza Yazid, Mastura Jaafar, Mustaffa Zainal, Aimi Salihah Abdul-Nasir and Noratikah Mazlan. "A review of image quality assessment (iqa): Snr, gcf, ad, nae, psnr, me." *Journal of advanced research in computing and applications* 7, no. 1 (2017): 1-7.

[23] Bera, Sutanu and Prabir Kumar Biswas. "Noise conscious training of non local neural network powered by self attentive spectral normalized Markovian patch GAN for low dose CT denoising." *IEEE Transactions on Medical Imaging* 40, no. 12 (2021): 3663-3673. https://doi.org/10.1109/TMI.2021.3094525

[24] Rusanov, Branimir, Ghulam Mubashar Hassan, Mark Reynolds, Mahsheed Sabet, Jake Kendrick, Pejman Rowshanfarzad and Martin Ebert. "Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: A systematic review." *Medical Physics* 49, no. 9 (2022): 6019-6054. https://doi.org/10.1002/mp.15840

[25] Shan, Hongming, Atul Padole, Fatemeh Homayounieh, Uwe Kruger, Ruhani Doda Khera, Chayanin Nitiwarangkul, Mannudeep K. Kalra and Ge Wang. "Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction." *Nature Machine Intelligence* 1, no. 6 (2019): 269-276. https://doi.org/10.1038/s42256-019-0057-9

[26] Gou, Shuiping, Wei Liu, Changzhe Jiao, Haofeng Liu, Yu Gu, Xiaopeng Zhang, Jin Lee and Licheng Jiao. "Gradient regularized convolutional neural networks for low-dose CT image enhancement." *Physics in Medicine & Biology* 64, no. 16 (2019): 165017. https://doi.org/10.1088/1361-6560/ab325e

[27] Gholizadeh-Ansari, Maryam, Javad Alirezaie and Paul Babyn. "Low-dose CT denoising using edge detection layer and perceptual loss." In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6247-6250. IEEE, 2019. https://doi.org/10.1109/EMBC.2019.8857940

[28] Liang, Tengfei, Yi Jin, Yidong Li and Tao Wang. "Edcnn: Edge enhancement-based densely connected network with compound loss for low-dose ct denoising." In *2020 15th IEEE International conference on signal processing (ICSP)*, vol. 1, pp. 193-198. IEEE, 2020. https://doi.org/10.1109/ICSP48669.2020.9320928

[29] Liu, Yan and Yi Zhang. "Low-dose CT restoration via stacked sparse denoising autoencoders." *Neurocomputing* 284 (2018): 80-89. https://doi.org/10.1016/j.neucom.2018.01.015

[30] Feng, Zhiwei, Ailong Cai, Yizhong Wang, Lei Li, Li Tong and Bin Yan. "Dual residual convolutional neural network (DRCNN) for low-dose CT imaging." *Journal of X-Ray Science and Technology* 29, no. 1 (2021): 91-109. https://doi.org/10.3233/XST-200777

[31] Feng, Zhiwei, Yizhong Wang and Ailong Cai. "Multi-Dimensional Spatial Attention Residual U-Net (Msaru-Net) for Low-Dose Lung Ct Image Restoration." In *Proceedings of the 4th International Conference on Advances in Image Processing*, pp. 55-60. 2020. https://doi.org/10.1145/3441250.3441266

[32] Zhang, Ju, Yan Niu, Zhibo Shangguan, Weiwei Gong and Yun Cheng. "A novel denoising method for CT images based on U-net and multi-attention." *Computers in Biology and Medicine* 152 (2023): 106387. https://doi.org/10.1016/j.compbiomed.2022.106387

[33] Jiao, Fengyuan, Zhiguo Gui, Yi Liu, Linhong Yao and Pengcheng Zhang. "Low-dose CT image denoising via frequency division and encoder-dual decoder GAN." *Signal, Image and Video Processing* 15, no. 8 (2021): 1907-1915. https://doi.org/10.1007/s11760-021-01935-0

[34] Park, Hyoung Suk, Jineon Baek, Sun Kyoung You, Jae Kyu Choi and Jin Keun Seo. "Unpaired image denoising using a generative adversarial network in X-ray CT." *IEEE Access* 7 (2019): 110414-110425. https://doi.org/10.1109/ACCESS.2019.2934178

[35] Chi, Jianning, Chengdong Wu, Xiaosheng Yu, Peng Ji and Hao Chu. "Single low-dose CT image denoising using a generative adversarial network with modified U-Net generator and multi-level discriminator." *IEEE Access* 8 (2020): 133470-133487. https://doi.org/10.1109/ACCESS.2020.3006512

[36] Suhaili, Shamsiah, Joyce Shing Yii Huong, Asrani Lit, Kuryati Kipli, Maimun Huja Husin, Mohamad Faizrizwan Mohd Sabri and Norhuzaimin Julai. "Development of Digital Image Processing Algorithms via FPGA Implementation." *Semarak International Journal of Electronic System Engineering* 3, no. 1 (2024): 28-45. https://doi.org/10.37934/sijese.3.1.2845

[37] Huang, Youyou, Rencheng Song, Kuiwen Xu, Xiuzhu Ye, Chang Li and Xun Chen. "Deep learning-based inverse scattering with structural similarity loss functions." *IEEE Sensors Journal* 21, no. 4 (2020): 4900-4907. https://doi.org/10.1109/JSEN.2020.3030321

[38] Hajewski, Jeffrey, Suely Oliveira and David Stewart. "Smoothed hinge loss and ℓ1 support vector machines." In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1217-1223. IEEE, 2018. https://doi.org/10.1109/ICDMW.2018.00174

[39] Jalil, Muhammad Al Imraan Abdul, Puteri Zirwatul Nadila Megat Zamanhuri, Syarmela Alaauldin and Asmalina Mohamed Saat. "Evaluation of Surface Crack Development on Marine Concrete Structure Using Matlab Image Processing." *Journal of Advanced Research in Computing and Applications* 21, no. 1 (2020): 1-11.

[40] Mohamed, Khaled H., Safa M. Gasser, Mohamed S. El-Mahallawy and Mohamed Waleed W. Fakhr. "Large scale image retrieval for remote sensing images using low level features." *Journal of Advanced Research in Computing and Applications* 8, no. 1 (2017): 8-14.

[41] Almaliki, Alaa Jabbar Qasim, Sajad Muhil Abd, Inam Abdullah Lafta, Roshidi Din, Osman Ghazali, Jabbar Qasim Almaliki and Sunariya Utama. "Application of the Canny Filter in Digital Steganography." *Journal of Advanced Research in Computing and Applications* 35, no. 1 (2024): 21-30. https://doi.org/10.37934/arca.35.1.2130