

Generative Adversarial framework for CT image denoising using Self-Attentive Residual UNet and Patch GAN Discriminator

Swapna Katta¹, Prabhishkek Singh^{2*}, Deepak Garg¹ and Manoj Diwakar³

¹School of Computer Science and Artificial Intelligence, SR University, Warangal, India.

²School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India.

³Department of CSE, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India.

*Corresponding author email: prabhisheksingh1988@gmail.com

<https://dx.doi.org/10.13005/bpj/3379>

(Received: 24 May 2025; accepted: 10 November 2025)

Computed Tomography (CT) is an indispensable tool to identify various health conditions. In Low Dose (LDCT) imaging, lower radiation is frequently chosen to minimize the impact of radiation on the human body, but it results in degraded image quality and subsequently generates noise in CT images. CT images are regularly influenced by Gaussian noise. The noise in CT images obscures fine anatomical details, thereby impacting accurate diagnostic precision. This is a challenging task for traditional image denoising approaches to maintain trade-off between reducing noise and preserving image information. Hence in LDCT images, lowering the radiation while enhancing the image quality is a critical challenge in clinical images. The main study explored the application of a GAN-based model that included the use of a self-attentive UNet generator to extract both local and global contextual data with a Patch GAN discriminator. The discriminator is evaluated to find the realism of local patches to maintain fine structural details. The quantitative measures such as the average Peak Signal to Noise Ratio (PSNR) value of 35.15 dB, and Structural Similarity Index Measure (SSIM) value of 0.92 at noise variance ($s = 10$) to examine the effectiveness of the model with respect to visual quality and clarity. The GAN-based model shows superior performance than ResNet50, UNet, and DnCNN models, showing that the GAN-based model is an optimistic denoising technique in LDCT images to suppress noise in LDCT images.

Keywords: Deep learning; GAN; Gaussian noise; Image denoising; LDCT imaging.

In recent years, CT has grown as an essential practical medical imaging technique for generating slice images of the inner anatomy to identify abnormalities such as cancer, fractures and tumors. CT imaging is a noninvasive technique that plays a multifaceted role in obtaining fine anatomical details using high temporal spatial resolution, and is effectively utilized to help pathological diagnosis and multiple treatment disciplines. However, the potential risks of X-ray

ionizing radiation have triggered public issue.¹ As per the known ALARA (As Low-as-Reasonably Achievable) standard, the reduction of X-radiation dose is a research hotspot in various medical imaging modalities, and the radiation dose might be reduced by minimizing X-ray intensity using criteria such as lessen the radiation exposure duration and tube current. The diminution of X-ray emission resulting noisy CT images. The noise in low dose CT imaging can be categorized as Gaussian and

Poisson noise.² Gaussian noise can occur due to the random variations in pixel intensity values and follows a Gaussian distribution, resulting in edge smoothness and uniform interference. Poisson noise can be identified by statistical behavior of X-ray photons, introduces mild yet noticeable grainy effect in the CT images. The random nature of pixel values generates noisy CT images that affects the accurate medical diagnosis. Therefore, it is necessary to minimize the noise in lower dose images by using different denoising methods.^{3,4}

Various image denoising techniques have been categorized as spatial domain filtering, transform domain filtering, and deep learning-based methods. Spatial filtering methods are applied directly to pixel values across a local neighborhood region. Spatial filtering methods like Gaussian, Wiener, and anisotropic diffusion, and total variation are effectively used to minimize noise and other minute artifacts.⁵⁻⁷ However, these methods often suffer from excessive noise smoothing while preserving the edges, leading to a loss of anatomical features that affect diagnostic precision. Among these, Total variation was used to solve sparse data issues in the CT images. In spatial domain filters, Block Matching 3D (BM3D) filtering and K-Singular Value Decomposition (KSVD) are popular methods. BM3D is a nonlocal, patch-based denoising method that is used in both spatial and transform domains, whereas KSVD is a dictionary learning and sparse-based adaptability and high-quality method for noise mitigation in lower dose images.^{8,9} Regardless of their strong performance, these techniques are often affected by the limitations of various kinds of noise or distortion in CT images, which exhibit nonstationary properties that can weaken their denoising performance. Transform domain filtering methods like wavelet transform, curvelet transform, and shearlet transform, decompose images into various frequency scales to suppress noise in various domains.^{10,11} Among these methods, the shearlet transform has superior anisotropic features and multidirectional capability to effectively handle noise and artifacts. However, improper thresholding can introduce ring and Gibbs artifacts around the edges of CT images.^{12,13}

Recently, Deep learning techniques have depicted superior outcomes in denoising LDCT imaging.^{14,15} DL-based techniques utilize neural

networks to identify complex noise patterns and correlations between image pixels, enabling them to adjust dynamically to varying noise conditions. DL-based techniques typically employ a training process that involves feeding a large dataset of noisy CT images and their associated noise-free counterparts to a neural network with objective of optimizing the network to diminish the variation between its predicted denoised and clean CT image.

Several deep learning models in particular CNN with different architectures, like Residual Encoder Decoder CNN (RED-CNN) and cascaded CNN, are integrated with different loss functions, such as MSE, Perceptual-loss, and adversarial-loss functions, and are effectively utilized during the learning stage of denoising model.^{16,17} These architectures are used to identify the complex nature of the denoising model, and loss functions, mainly recognize how effectively model learns from related data. In general, GAN-based denoising models can generate superior results by effectively utilizing distribution transformation and overcoming the voxel-level regression MSE, which can effectively mitigate the over smoothing of LDCT images. Hybrid denoising methods are fundamental to CT image denoising, as human vision perceived quality does not depend on pixel-by-pixel variations but is instead sensitive to contextual and structural features. To overcome this, WGAN-based noise suppression was integrated with a pretrained VGG perceptual loss function. The VGG network is trained on standard RGB images, the use of its perceptual loss function in LDCT images often generates cross-hatch artifacts, and a degradation of crucial details.¹⁸

A novel hybrid GAN method optimizes LDCT image quality using a Generator network to convert noisy images into clean images, whereas the discriminator distinguishes between real and synthetic (generated) outputs. This adversarial approach enhances imaging quality and contributes to higher diagnostic precision.¹⁹ The hybrid approach DU-GAN effectively utilizes GANs and UNet-based discriminators to enhance imaging quality. However, this model is computationally intensive.²⁰ Progressive Wasserstein Generative Adversarial Network (PWGAN) is the novel hybrid approach, including a weighted structural sensitivity hybrid loss function to enhance noise suppression in LDCT images.²¹ The hybrid Dual

Encoder Single Decoder (DESD) is a combination of a generator, and pyramid nonlocal attention module to perform feature correlation and auxiliary shallow and deep feature process modules to improve feature extraction. However, this model is highly sensitive to hyperparameters.²² Artifact and detail attention GAN utilizes a multichannel generator to suppress noise, other artifacts, edge details, and multiscale discriminators to enhance discriminative power. However, the computational cost is high.²³

Related work

Recently, GAN-based techniques have become a strong framework for LDCT image denoising, providing superior performance in maintaining edge details and structural information compared with conventional deep learning methods. This portion reviews various deep learning-based approaches, and existing GAN related approaches for noise suppression and image enhancement in LDCT images. The Saidulu *et al.*²⁴ proposed Asymmetric Convolution-based GAN framework (ACG-Net), which is a GAN-based LDCT image denoising approach that includes 1D-Asymmetric convolutions to improve directional features and dynamic attention to fuse multiscale data to maintain both local and global pixel correlations, as well as Neural Structure Preserving Loss (NSPL) to maintain the integrity of anatomical structures. The method provides enhanced structural fidelity and higher perceptual quality to improve denoising performance in LDCT images. However, the model is complex, and demanding computational requirements may inhibit its significance in real time medical situations. Kuraning *et al.*²⁵ proposed a CycleGAN-based LDCT denoising technique that integrates a UNet Generator for detail preservation and Patch-GAN discriminator for texture realism to reduce noise in CT images. This method efficiently handles noise mitigation and detail preservation to enhance the overall imaging quality, thereby improving diagnostic precision. However, this method is computationally intensive, which may lead to difficulties in real-time clinical settings.

Abuya *et al.*²⁶ introduced a novel hybrid approach, termed Adversarial Content-Noise Complementary Learning (ACNCL), which combines multiple noise reduction approaches such as DnCNN, UNet, CA-AGF, and Discrete Wavelet

transform (DWT) within a GAN framework. The method further utilizes a patch-GAN discriminator to preserve fine image features and thereby improve overall visual quality. The ACNCL distinguishes noise and anatomical features to ensure effective noise minimization while maintaining structural information. The method showed superior denoising performance on CT and MRI datasets to improve image clarity for facilitating prior tumor detection. However, this method incurs computational overhead owing to the use of multiple denoising approaches. Lokhande *et al.*²⁷ implemented a deep dense GAN is integrated with loss function for unsupervised blind noise mitigation, efficiently handling diverse noise types without using paired noisy-clean image datasets. This method effectively utilizes deep and Generative GAN with a novel loss function to suppress noise at various intensity levels. The novel method provides effective noise reduction, fine detail preservation, robust performance in handling complex noise patterns, and artifact removal, thereby facilitating better diagnostic precision.

He *et al.*²⁸ introduced the Structural Semantic Enhancement Network (SSEN) to improve the image sharpness by effectively utilizing a 5x5 sobel operator with dense connections and asymmetric convolutional layers to enhance structural-feature extraction. It preserves both fine image details and structural patterns more efficiently. The compound loss function, combining L1 loss, Multiscale SSIM loss, and multiscale perceptual-loss, was carried out to restore structural and perceptual integrity, yielding better PSNR and SSIM performance. However, the advanced model structure incurs higher computational costs and requires careful parameter tuning. Wang *et al.*²⁹ assessed the various advanced GAN-based techniques for suppressing noise while conserving fine details in LDCT images. This review outlines the main architectural enhancements, such as conditional GANs (C-GANs), Super-resolution GANs (SRGANs), and Cycle-GANs, and Quantitative evaluation of various denoising models using PSNR, SSIM, and Learned-Perceptual-Image-Patch-Similarity (LPIPS) metrics. The GAN architectures show superior denoising, but limitations might exist in reliability, and synthetic artifact generation still exists.

Jiao *et al.*³⁰ implemented a frequency division GAN integrated with an encoder and dual decoder architecture, which provides noise suppression by implementing the isolation of high-frequency coefficients for effective denoising while preserving tissue edge details and textures. The dual decoder recovers lost details, while the denoising process and multiscale inception discriminator improve feature-level discrimination. This method achieved superior noise reduction. However, a complex architectural design might result in higher computational costs and increased training times. Wang *et al.*³¹ suggested that a novel Residual Structure and Cooperative Attention framework is based on GAN (RCA-GAN). This method incorporates a residual learning and cooperative attention module, integrated with a multimodal loss function to enhance the effectiveness of the noise reduction process. The model demonstrated superior noise suppression and fine detail preservation. However, this method is also computationally intensive.

Recently, Naser *et al.*³² proposed a Denoising GAN-based network for CT image denoising, incorporating a Recursive Residual Group-based generator network with Selective Kernel-Feature-Fusion (SKFF), and a Patch GAN discriminator. This network effectively enhances feature refinement, texture modelling, and hierarchical context-capture in Cardiac Magnetic Resonance imaging (CMR). This model achieved higher PSNR and SSIM values across varying noise intensities, demonstrating the performance of RRG method to suppress noise while preserving fine image details. While their model proved effective for CMR images, our study focuses on LDCT image denoising, which presents distinct challenges due to different noise characteristics, anatomical structures, and diagnostic constraints. To address these constraints, we leverage a self-attention based UNet generator and a Patch GAN discriminator, optimized with adversarial and L1 fidelity loss, to ensure preservation of anatomical details and diagnostic fidelity in LDCT images.

MATERIALS AND METHODS

GAN-based methodology

The GAN-based denoising technique supports paired images such as noisy and clean

images, and converts noisy CT images into denoised ones. Fig.1 depicts the generalized methodology used in image denoising to improve image quality.³³The GAN-based technique was initially applied to CT images corrupted with an Additive Gaussian noise. The Gaussian noise exhibits a normal distribution having a mean value of zero and a standard deviation, is widely indicated as a bell-shaped curve.

Mathematically, it is represented as:

$$I_j^*(a,b) = I_j(a,b) + N_j(a,b) \quad \dots(1)$$

$I_j^*(a,b)$ is the noisy CT image

$I_j(a,b)$ is the clean CT image

$N_j(a,b)$ is the Additive Gaussian noise

(a, b) represents the pixel positions in the CT image

To demonstrate the denoising efficiency of the GAN-based model for LDCT imaging, a “Large covid-19 CT-Scan slice Dataset” was applied to measure the quality of the denoised CT images.³⁴ This dataset comprises clean CT images collected from different patients, each image with a spatial resolution of 256 x 256 pixels. A total of 7593 CT images, of which 6074 are allocated for training and 1519 are intended for testing. The model was fed with a pair of clean and Gaussian noisy images. A comparison between different denoising methods, such as ResNet50³⁵, UNet³⁶, and DnCNN³⁷, with a GAN-based model to identify the denoising performance of different methods to improve overall imaging quality.

The GAN-based architecture for CT image denoising integrates the following features:

A residual UNet Generator with self-attention, a patch GAN discriminator, and Adversarial + L1 loss function are used to improve the image quality. Quantitative benchmarks were used to examine the denoising imaging quality using PSNR and SSIM.

Input:

Initially take a dataset with a set of paired CT images

$$D_s = \{(\tilde{I}_j, I_j)\}_{j=1}^N \quad \dots(2)$$

D_s : Dataset

$\tilde{I}_j \in \mathbb{R}^{h \times w}$: Gaussian Noisy CT image

$I_j \in \mathbb{R}^{h \times w}$: Clean Ground truth CT image

h : Number of rows

w : Number of columns

Define Loss functions:

Binary Cross entropy (BCE) loss-function to get adversarial loss

L1 loss function for reconstruction

Step 1: Implement a Generator (G)

A deep CNN including a UNet with residual and self-attention blocks, which can learn the mapping, it can be expressed as:

$$G : \tilde{I}_j \rightarrow \hat{I}_j = G(\tilde{I}_j) \quad \dots(3)$$

\hat{I}_j : Denoised CT images generated from \tilde{I}_j .

Step 2: Implement a Patch-GAN Discriminator (D)

A Patch GAN Discriminator is a neural network, trained to discriminate between real/fake CT images, and is expressed as:

$$D : (\tilde{I}_j, I_j) \rightarrow [0,1], \text{ and } D : (\tilde{I}_j, \hat{I}_j) \rightarrow [0,1] \quad \dots(4)$$

The Discriminator network learns to compare a pair of (noisy, clean) CT images and provides an output value in a specific range between 0 and 1.

If the second CT image in the pair is (ground truth) clean CT image, the resulting output is 1, and is real pair of CT images (\tilde{I}_j, I_j) .

If the second CT image in the pair is generated by the Generator, resulting output is 0, represents fake generated pair of CT images: (\tilde{I}_j, \hat{I}_j)

Step 3: Apply Loss functions

a). Calculate Adversarial loss for Generator

The Generator network attempts to believe (fool) that results in a discriminator output \tilde{I}_j that is closer to the real (ground-truth) clean image I_j .

In general, loss function measures the performance of the Generator (G) in a GAN.

The ideal GAN loss function expects the Discriminator to generate a high probability

(approximately 1) of generated CT images.

The loss function can be calculated as:

$$\mathcal{L}_{GAN}(G) = \mathbb{E}_{\tilde{I}_j} [\log D(\tilde{I}_j, G(\tilde{I}_j))] \quad \dots(5)$$

$G(\tilde{I}_j) = \hat{I}_j$: Generated denoised CT image

$D(\tilde{I}_j, G(\tilde{I}_j))$: Discriminator's confidence that the generated CT image is real

The Generator network was trained to maximize its output quality, enabling the Discriminator to classify the generated CT images as real.

In practice, we preferred Binary-Cross-Entropy loss is utilized for numerical stability

$$\mathcal{L}_{GAN}(G) = \mathbb{E}_{\tilde{I}_j} [BCE(D(\tilde{I}_j, G(\tilde{I}_j)), 1)] \quad \dots(6)$$

Where:

BCE : Binary Cross Entropy loss

$D(\tilde{I}_j, G(\tilde{I}_j))$: Discriminator-generated output.

The desired output was 1, ensuring that the discriminator was misled to believe that the generated CT image was real.

Pixel-wise reconstruction loss (L1 loss) function:

In addition to adversarial loss, a pixel level fidelity between the generated denoised image and ground-truth clean CT image is enforced using the L1 reconstruction loss function, which can be denoted as:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{(\tilde{I}_j, I_j)} [\|G(\tilde{I}_j) - I_j\|_1] \quad \dots(7)$$

Scaling can be performed by weight ($\lambda = 100$) is empirically chosen for reconstruction loss:

$$\mathcal{L}_{recon}(G) = \lambda \cdot \mathcal{L}_{L1}(G) \quad \dots(8)$$

λ : The Hyperparameter (weighting factor) controls the significance of the L1 loss function during training process.

$\mathcal{L}_{recon}(G)$: Reconstruction loss represents how accurately Generator network (G) is reconstructing the clean CT image from a noisy image.

Total Generator loss function:

The total generator loss function is used to combine Adversarial loss \mathcal{L}_{GAN} and Reconstruction loss \mathcal{L}_{L1} , which can be expressed as:

$$\mathcal{L}_G = \mathcal{L}_{GAN}(G) + \lambda \cdot \mathcal{L}_{L1}(G) \quad \dots(9)$$

Update Generator network parameters using Adam optimizer:

$$\phi_G = Adam(\nabla_{\phi_G} \mathcal{L}_G, \delta G, \xi_1 = 0.5, \xi_2 = 0.99) \quad \dots(10)$$

learning rate (δG) = 2×10^{-4}

$\xi_1 = 0.5$, $\xi_2 = 0.99$ are hyperparameters are used to stabilize the learning process.

$\nabla_{\phi_G} \mathcal{L}_G$ is serves to calculate the gradient of the Generator loss (\mathcal{L}_G) with respective to the Generator network parameters (ϕ_G)

Discriminator loss function:

The Discriminator network is trained to accurately categorize real vs. fake images and can be expressed as:

$$\mathcal{L}_D = \frac{1}{2} \mathbb{E}_{(\tilde{I}_j, I_j)} [BCE(D(\tilde{I}_j, I_j), 1)] + \frac{1}{2} \mathbb{E}_{\tilde{I}_j} [BCE(D(\tilde{I}_j, G(\tilde{I}_j)), 0)] \quad \dots(11)$$

(\tilde{I}_j, I_j) : Real pair of CT images, I_j represents the clean image

$(\tilde{I}_j, G(\tilde{I}_j))$: Fake pair of CT images, where $G(\tilde{I}_j)$ represents the denoised output from the Generator.

The Discriminator loss can be defined as the average of the two binary cross entropy losses, that is, one indicated as correctly classifying real pairs of images as real and another to classify the fake pair of CT images as fake.

Update discriminator network parameters using the Adam optimizer:

$$\phi_D = Adam(\nabla_{\phi_D} \mathcal{L}_D, \delta D, \xi_1 = 0.5, \xi_2 = 0.99) \quad \dots(12)$$

Learning rate: $\delta D = 2 \times 10^{-4}$

$\nabla_{\phi_D} \mathcal{L}_D$: is used to calculate the gradient of the Generator loss (\mathcal{L}_D) with respective to the Discriminator network parameters ϕ_D .

Output:

Display denoised CT images.

Evaluate quantitative metrics like (PSNR and SSIM), and Generator and Discriminator losses

$(\mathcal{L}_G, \mathcal{L}_D)$ to evaluate CT imaging quality.

Algorithm explanation

GAN-based denoising framework employs a Generator and Discriminator architecture for denoising LDCT images, aiming to optimize structural integrity and preserve fine details. The set of noisy (\tilde{I}_j) and clean (I_j) CT image pairs from a dataset $D = \{(\tilde{I}_j, I_j)\}_{j=1}^N$, which can be used to train the GAN-based framework with a predefined number of epochs. The Generator was implemented as a Residual UNet integrated with a self-attention model to extract both local textures and global contextual information that is significant in medical CT imaging to suppress noise in LDCT images.

For a given Gaussian noisy image \tilde{I}_j , the Generator (G) generates a denoised estimate (\hat{I}_j) = $G(\tilde{I}_j)$. The Discriminator (D) is implemented based on the patch GAN architecture to evaluate the realism of coinciding local patches instead of the whole image. This architecture allows the discriminator to emphasize the fine-grained texture details and structural patterns to guide the Generator framework to maintain detailed anatomical information. The Generator can be optimized using a composite loss function, which comprises two provisions such as Adversarial loss ($\mathcal{L}_{GAN}(G)$) and L1 Reconstruction loss ($\mathcal{L}_{L1}(G)$). An Adversarial loss function facilitates the Generator to generate images that are identical from ground-truth clean CT images, which can be evaluated by the Discriminator (D). The L1 Reconstruction loss can be calculated as the mean absolute error between the generated CT image and ground truth clean image. This function implements pixel-level accuracy and effectively maintains edge sharpness and fine structural information. The total generator loss was calculated as follows:

$$\mathcal{L}_G = \mathcal{L}_{GAN}(G) + \lambda \cdot \mathcal{L}_{L1}(G) \quad \dots(13)$$

Where λ denotes a weighting factor to balance the perceptual realism and pixel-level fidelity. In our implementation, we set $\lambda=100$ to focus on preserving the structural details while maintaining a realistic reconstruction accuracy.

To obtain stable adversarial training and efficient convergence of the Generator and Discriminator, different learning rates are applied to each optimizer. An Adam optimizer can be used

to optimize the model parameters with learning rates ($\delta G = 2 \times 10^{-4}$, $\delta D = 2 \times 10^{-4}$) for the Generator and Discriminator accordingly. The noisy and clean CT image pairs are given as input to the network. An Adam optimizer can be used for Generator and Discriminator, with the following hyperparameters.

The Generator and Discriminator are updated using an Adam optimizer to minimize corresponding loss functions.

RESULTS

Quantitative Analysis

Quantitative metrics, such as the PSNR and SSIM, were applied to the test set to examine the denoising performance of the model.^{38,39} Additionally, denoised images were quantitatively assessed and visually compared with the respective noisy-clean reference images.

Peak Signal to Noise Ratio:

It was used to measure the performance of the denoising process during image reconstruction. It assesses the denoised image quality using the original image quality.

$$PSNR(I_j, \hat{I}_j) = 10 \cdot \log_{10} \left(\frac{255 \times 255}{MSE(I_j, \hat{I}_j)} \right) \quad \dots(14)$$

Where:

I_j : is the clean image

\hat{I}_j : is the denoised image

Mean_squared_error (MSE): This is denoted as the average square difference between the original CT image and denoised one.

Structural Similarity Index Measure (SSIM):

This is a perceptual metric for examining the similarity within two images. SSIM is used to identify the variations in structural information,

contrast and luminance to measure the visual structure preservation between the two images.

$$SSIM(I_j, \hat{I}_j) = \frac{(2\mu_{I_j}\mu_{\hat{I}_j} + c_1)(2\sigma_{I_j\hat{I}_j} + c_2)}{(\mu_{I_j}^2 + \mu_{\hat{I}_j}^2 + c_1)(\sigma_{I_j}^2 + \sigma_{\hat{I}_j}^2 + c_2)} \quad \dots(15)$$

Where:

$\mu_{I_j}, \mu_{\hat{I}_j}$: Mean of I_j and \hat{I}_j images

$\sigma_{I_j}^2, \sigma_{\hat{I}_j}^2$: Variance of I_j and \hat{I}_j

$\sigma_{I_j \hat{I}_j}$: Covariance between I_j and \hat{I}_j

C_1, C_2 : Small constant values to stabilize division ($C_1 = (0.01 \cdot L)^2$, $C_2 = (0.03 \cdot L)^2$, and L is the dynamic range, mainly 255 for 8-bit CT images)

The range of SSIM is $[-1, 1]$, in practice generally $[0, 1]$, where 1 indicates perfect similarity and 0 means there is no similarity between two images.

Comparison of results

Table 1 depicts the average PSNR and SSIM values for different denoising methods, such as ResNet50³⁵, UNet³⁶, DnCNN³⁷ and GAN-based methods, which were evaluated using two different noise intensity levels ($\delta=10$ and $\delta=20$), and graphical notations of PSNR and SSIM are shown in Figures 2 and 3. The experimental outcomes depict that the GAN-based model consistently outperforms other denoising methods in both the quantitative evaluation metrics. For Gaussian noise intensity ($\delta=10$), the GAN-based model reached an average PSNR value of 35.15 Db, and an SSIM value of 0.89, showing superior denoising while maintaining structural information. In comparison with ResNet50³⁵ yielded lower denoising performance with a PSNR value is of 31.21 dB, indicating its limited noise-suppression ability to restore finer textures.

The denoising performance of the methods was slightly reduced for noise intensities ($\delta=20$) owing to increased noise complexity.

Table 1. The average PSNR and SSIM values for different denoising models

Denoising method	Avg. PSNR		Avg. SSIM	
	Noise variance (σ) 10	Noise variance (σ) 20	Noise variance(σ) 10	Noise variance(σ) 20
ResNet50 ³⁵	31.21	30.20	0.89	0.85
UNet ³⁶	32.45	31.50	0.91	0.90
DnCNN ³⁷	34.16	32.10	0.91	0.89
GAN-based method	35.15	34.53	0.92	0.91

However, the GAN-based model has robustness with a PSNR value of 34.53 dB and SSIM value of 0.91. The DnCNN³⁷, which has obtained PSNR 32.1 dB and SSIM value 0.89 respectively and UNet³⁶ has competitive SSIM values and lower PSNR value, indicating that it preserves structural details well but not mitigate noise effectively. Overall, the GAN-based model shows best trade-off between noise mitigation and fine detail conservation, resulting in a strong method for CT image denoising to maintain image quality, and fine anatomical details. Fig.4 shows the denoising performance of epochs (5,10,15 and 20) for various denoising methods. The GAN-based model consistently shows the highest PSNR, and Fig. 5 depicts the highest SSIM, indicating superior denoising performance while preserving structural fidelity compared with ResNet50³⁵, UNet³⁶, and DnCNN.³⁷ The Generator and Discriminator losses per each epoch of GAN-based denoising approach are shown in Fig. 6. The graph depicts the training process of the GAN using a Generator

and Discriminator loss over 20 epochs. In the Initial stage, the Generator loss increased, peaking at approximately epochs 10-12, which signifies that the Generator struggles to produce convincing results as the Discriminator efficiently recognize fake samples. Subsequently, the Generator loss decreases, indicating an improvement in the generation of realistic images. The discriminator loss was stable and consistently low during training, indicating that it had a strong ability to distinguish in between real and fake data.

Time complexity

Visual Quality Analysis

A visual comparison of the denoised Gaussian noisy images at noise variance 10 of four different CT images with the GAN-based model is shown in Fig. 7,8,9, and 10. The Figures indicate that the denoised outcomes have better visual quality, indicating the better denoising ability of the GAN-based model. A visual assessment of denoising results from different denoising techniques such as ResNet50³⁵, UNet³⁶, DnCNN³⁷

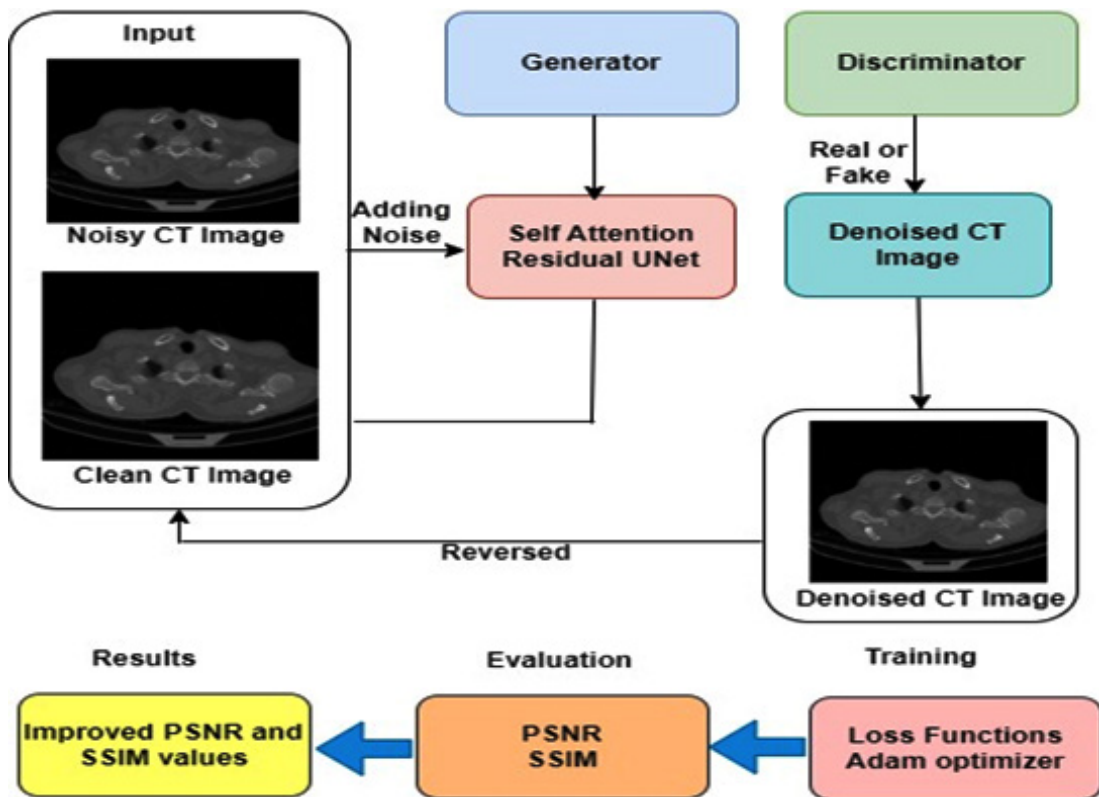


Fig. 1. General GAN-based denoising methodology.

and GAN-based model of two sample images is depicted in Figures 11 and 12. ResNet50³⁵ depicts moderate denoising ability owing to the over-smoothness of fine structural information and might struggle to capture broader contextual information of CT images. UNet³⁶ showed that the effective visual quality of CT images can result in blurriness in the edge details. The DnCNN³⁷ depicts better denoising of Gaussian noisy images. However the model struggles to maintain fine textures, which significantly affects the visual quality. After

implementing the GAN-based model on noisy CT images, it was observed that the denoised images had superior visual quality compared with other LDCT denoising models.

DISCUSSION

This study demonstrated a self-attention GAN-based architecture for suppressing Gaussian noise in LDCT images. The Generator network integrates a UNet architecture with residual

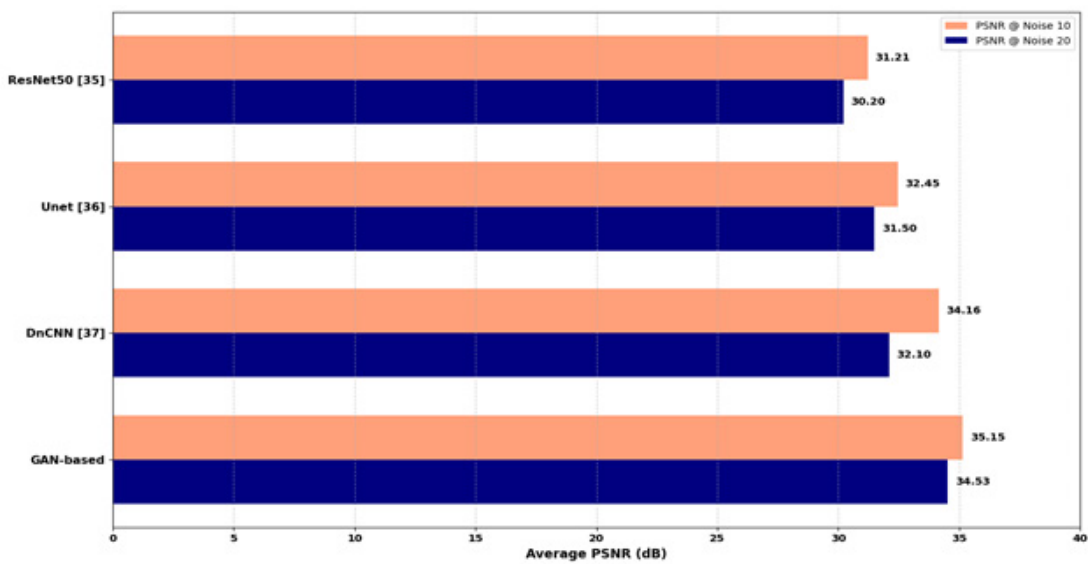


Fig. 2. Graphical notation of Average PSNR at Gaussian noise variance ($\sigma = 10$ and 20)

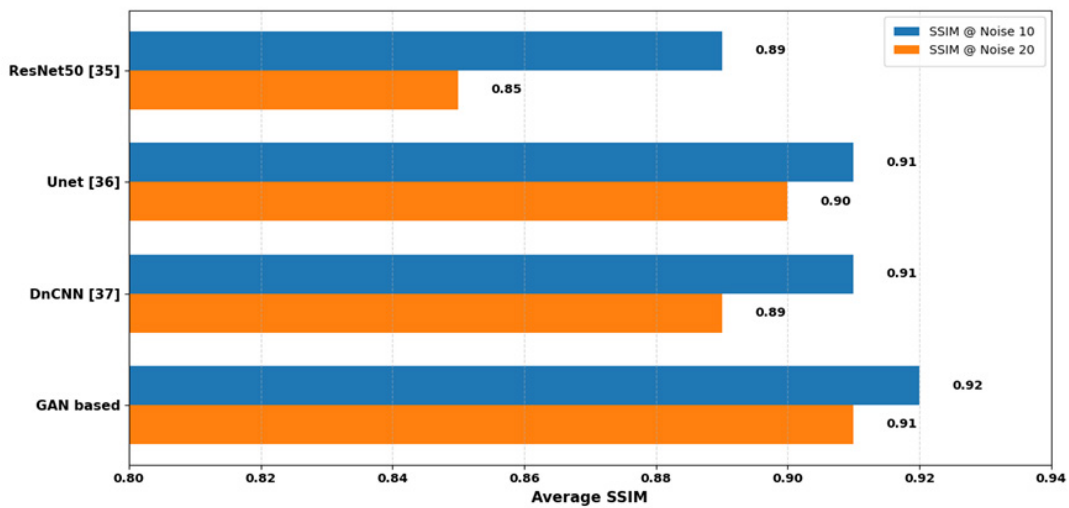


Fig. 3. Graphical notation of Average SSIM at Gaussian noise variance ($\sigma = 10$ and 20)

connections and a self-attention mechanism. In contrast, a Patch GAN discriminator network is leveraged to evaluate the originality of the denoised output images. The self-attention mechanism plays a major role in capturing long-range contextual dependencies to improve the quality of LDCT images. The training strategy combines the adversarial loss and L1 reconstruction loss, enabling the network to maintain fine

structural details while effectively mitigating noise. Quantitative results show that the GAN-based model achieved the highest average PSNR and SSIM values, with a PSNR of 35.15 dB at Gaussian noise level ($\sigma=10$) and 34.53 dB at ($\sigma=20$), and the corresponding average SSIM values were 0.92 and 0.91, respectively, as illustrated in Table 1. The experimental results indicate that the GAN-based model outperformed existing denoising

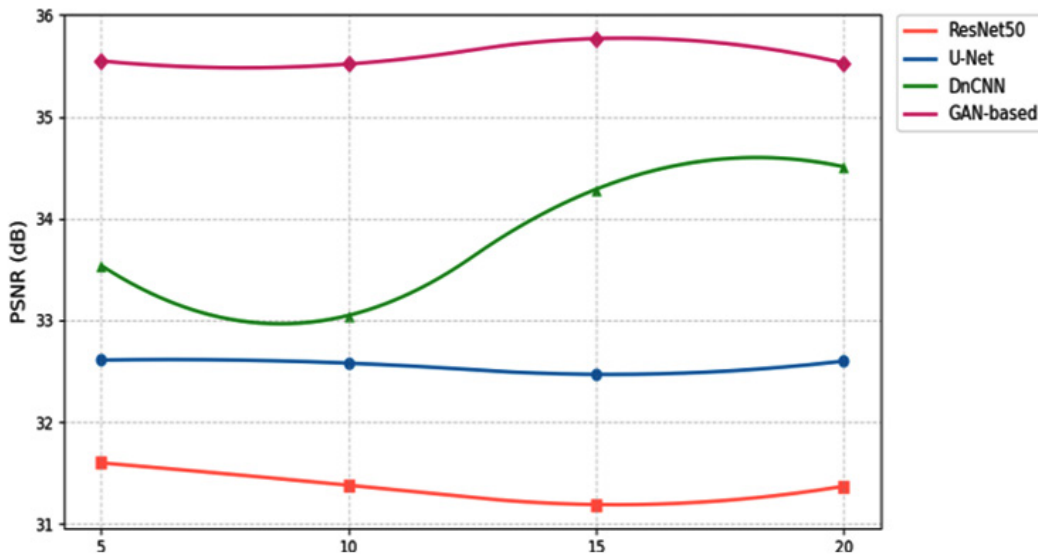


Fig. 4. PSNR comparison at Selected epochs (5,10,15,20) with Gaussian noise variance ($\sigma=10$)

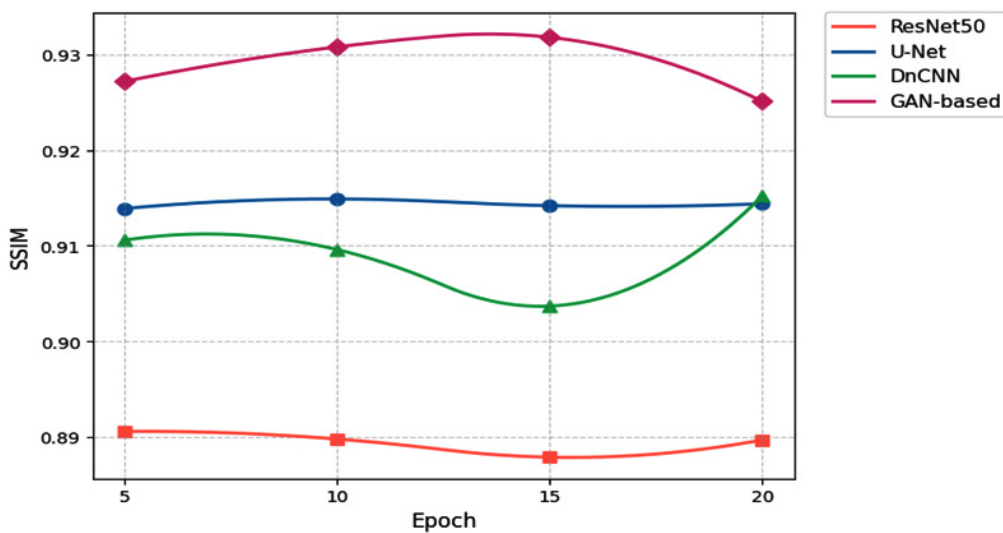


Fig. 5. SSIM comparison at Selected epochs (5,10,15,20) with Gaussian noise variance ($\sigma=10$)

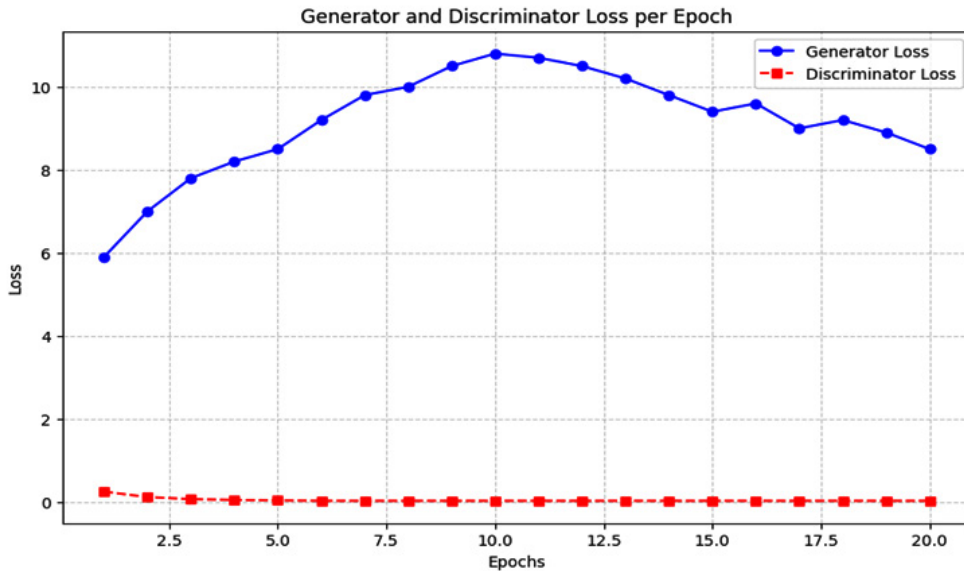


Fig. 6. Generator and Discriminator loss per each epoch of GAN-based denoising model

architectures, such as ResNet50³⁵, UNet³⁶, and DnCNN³⁷, with respect to PSNR and SSIM values, ensuring high fidelity in the reconstructed CT images. The ResNet50³⁵ and UNet³⁶ denoising models resulted in over-smoothing, and the DnCNN³⁷ architecture struggled to maintain fine anatomical details at higher noise levels. Visual analysis of the GAN-based model provided better visual insights into the denoised CT images, highlighting the model’s effectiveness in clinical lower dose CT applications. Overall, the experimental findings validate that the GAN-based framework is superior in suppressing Gaussian noise while preserving structural details.

Table 2 demonstrates a detailed comparative evaluation of computational cost and Training time-complexity of various CT image denoising methods including ResNet50³⁵, UNet³⁶, DnCNN³⁷, and GAN-based model (UNet+ Patch-Discriminator with self-attention), emphasizing their balance between effectiveness and modeling capability. Each batch wise cost (\tilde{F}_{model}) is computed using CNN kernel size (\tilde{K}), number of input/output (C_{in}, C_{out}) channels, and architectural specific models like Upsampling with convolution (UpConv), Deconvolution (DeConv), and self-attention layers. The total model’s Training time complexity is defined in

Big(O) notation, and is denoted as $O(\tilde{E} \cdot \frac{N}{B} \cdot \tilde{F}_{model})$. DnCNN-based denoising method has 17 CNN layers. The computational cost is based on three aspects: The first aspect is the number of network weights associated with each CNN layer, denoted by \tilde{K}^2 . The second aspect is the evaluation of the expression, $\sum_{h=1}^{17} \tilde{K}^2 \times C_{in,h} \times C_{out,h}$, where the summation is performed over all layers. The third aspect is the total processing cost of all batches, represented by $O(\tilde{E} \cdot \frac{N}{B} \cdot \sum_{h=1}^{17} \tilde{K}^2 C_{in,h} C_{out,h})$. This method results in computationally light, fixed in image resolution, but is often less capable of extracting contextual information. UNet-based denoising method has an encoder to capture features from a noisy CT image, and a decoder network to reconstruct the clean CT image from extracted features, incorporating skip connections, a bottleneck computation, raising complexity but improving multiscale feature extraction. ResNet50-based denoising method has a higher cost due to its deep-encoder network with additional layers and channels. The GAN model is computationally challenging, with a $\tilde{F}_{G,conv} + \tilde{S}_a^2 \cdot C_a + \tilde{F}_{D,conv}$, whereas $\tilde{S}_a^2 \cdot C_a$ signifies quadratic cost of self-

Table 2. Theoretical Time Complexity Comparison of various denoising methods

Denoising Method	Batch wise cost $\tilde{F}_{_model}$	Training time complexity Big O (.)
ResNet50 ³⁵	ResNet50 encoder: $\sum_{h \in ResNet50} \hat{K}^2 C_{in,h} C_{out,h}$ ResNet50 decoder: $\sum_{up} DeConv + Conv$	$\tilde{E} \cdot \frac{\mathcal{N}}{\bar{B}} \cdot \tilde{F}_{ResNet50den}$
UNet ³⁶	Encoder: $\sum_{enc} 2 \hat{K}^2 C_{in} C_{out}$ Bottleneck: $2 \hat{K}^2 C_{in} C_{out}$ Decoder: $\sum_{dec} [UpConv + 2 \hat{K}^2 C_{in} C_{out}]$	$\tilde{E} \cdot \frac{\mathcal{N}}{\bar{B}} \cdot \tilde{F}_{UNet}$
DnCNN ³⁷	$\sum_{h=1}^{17} \hat{K}^2 C_{in,h} C_{out,h}$	$\tilde{E} \cdot \frac{\mathcal{N}}{\bar{B}} \cdot \sum_{h=1}^{17} \hat{K}^2 C_{in,h} C_{out,h}$
GAN (UNet-Generator, Patch-based Discriminator +Self attention)	$\tilde{F}_{G,conv} + \tilde{F}_{G,attn} + \tilde{F}_{D,conv}$ Where: $\tilde{F}_{G,attn} = \check{S}_a^2 \cdot C_a$	$\tilde{E} \cdot \frac{\mathcal{N}}{\bar{B}} \cdot (\tilde{F}_{G,conv} + \check{S}_a^2 C_a + F_{D,conv})$

Where:

Training variables

\tilde{E} : Number of Epochs

\mathcal{N} : Number of training sample images

\bar{B} : Batch size

Convolution-layer parameters

\hat{K} : Size of the kernel

$C_{in,h}$: Number of (input) channels for layer h

$C_{out,h}$: Number of (output) channels for layer h

$\sum_{h=1}^{17}$: Summation of Convolution layers

Architecture Specific Notations

Encoder/Decoder: Specific parts of network for encoder(downsample) or decoder for upsample the feature mappings

UpConv : Upsampling with convolution

DeConv : Deconvolution or Transposed Convolution

Bottleneck : Middle-block in between encoder/decoder

$\tilde{F}_{ResNet50den}$: Computational cost of each batch of the ResNet50 denoising model

\tilde{F}_{UNet} : Computational cost of each batch of the UNet denoising model

GAN specific Notations

$\tilde{F}_{G,conv}$: Convolution layer computational cost of the GAN network generator

$\tilde{F}_{D,conv}$: Convolution layer computational cost of the GAN network discriminator

$\tilde{F}_{G,attn}$: Attention-based computational cost of the Generator

\check{S}_a^2 : Spatial dimensions (height/width) of the feature map

C_a : Total number of attention channels

attention all over spatial size \bar{s}_a and attention-channels (C_a). The higher complexity of this model is well substantiated by GAN's superior capability to suppress noise, extract contextual dependencies while preserving fine structural

details, and generate perceptually intensive and high-quality CT images.

Visual quality of denoised CT images can be analyzed using an intensity profile. It is mainly used to visualize image structure in terms of how image brightness varies, showing edges, and fine

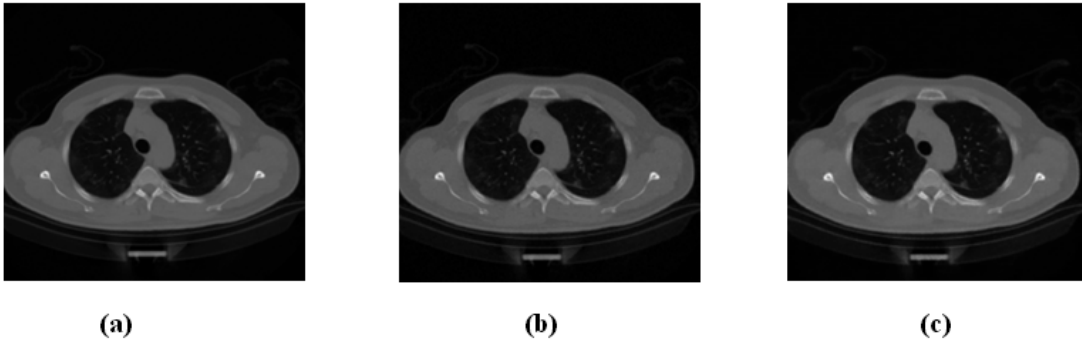


Fig. 7. (a) Ground truth CT1 Image; (b) Noisy CT1 Image at noise variance ($\sigma=10$); (c) GAN-based model Denoised CT1 Image.

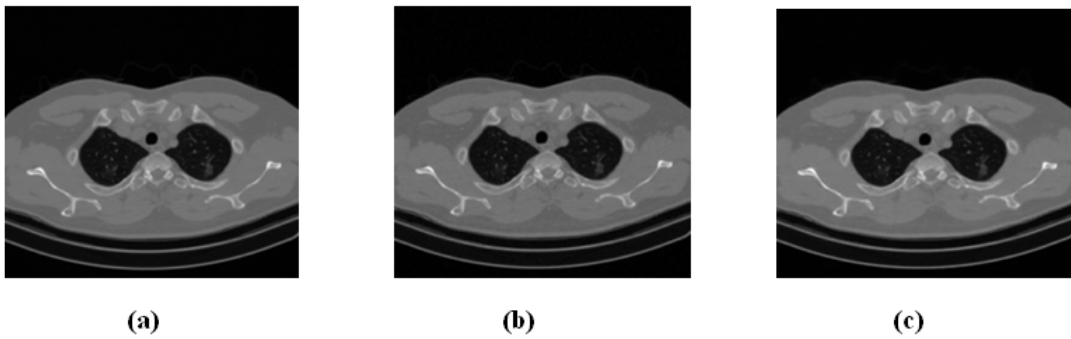


Fig. 8. (a) Ground truth CT2 Image; (b) Noisy CT2 image at noise variance ($\sigma=10$); (c) GAN-based model Denoised CT2 Image.

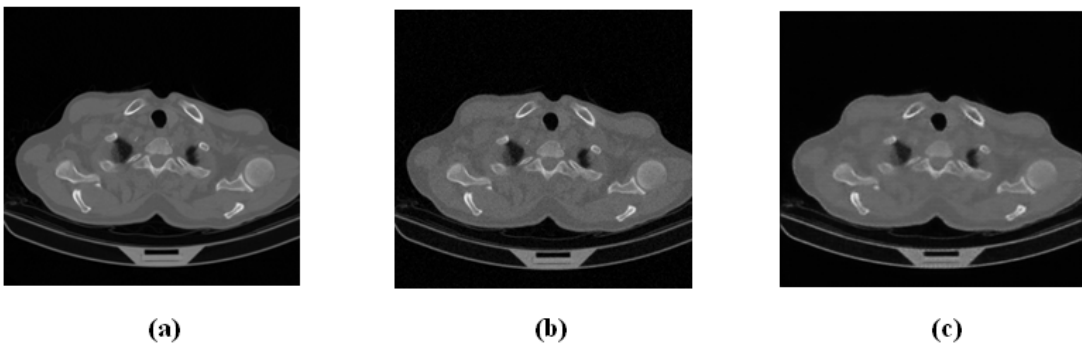


Fig. 9. (a) Ground truth CT3 Image; (b) Noisy CT3 Image at noise variance ($\sigma=10$); (c) GAN-based model Denoised CT3 Image.

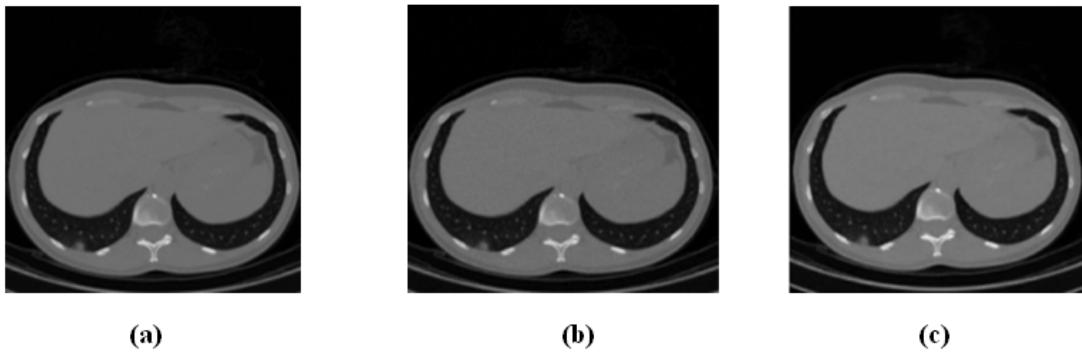


Fig. 10. (a) Ground truth CT4 Image; (b) Noisy CT4 Image at noise variance ($\sigma=10$); (c) GAN-based model Denoised CT4 Image.

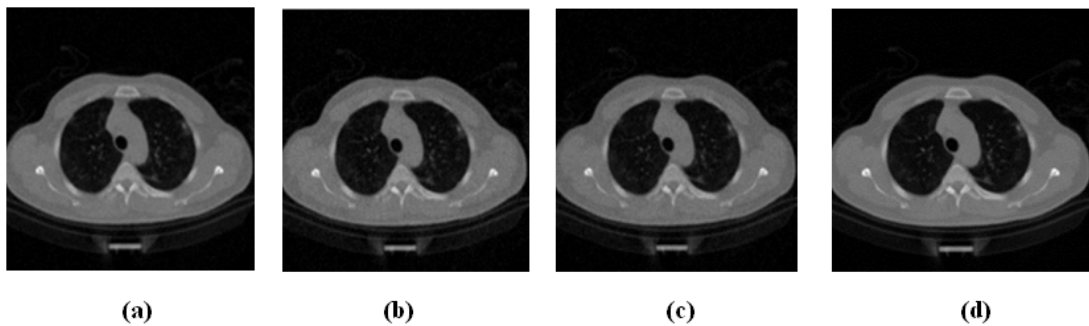


Fig. 11. Denoising outcomes of different denoising models of CT1 Image; (a) ResNet50 [35]; (b) UNet [36]; (c) DnCNN [37]; GAN-based model.

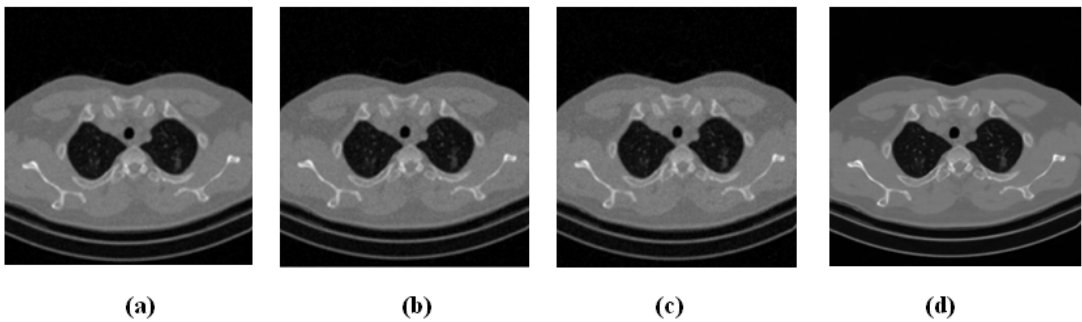


Fig. 12. Denoising results of different denoising models of CT2 Image; (a) ResNet50 [35]; (b) UNet [36]; (c) DnCNN [37]; (d) GAN-based model.

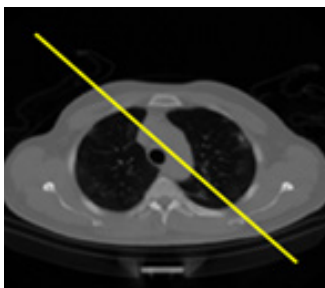


Fig. 13. Line segment of Clean CT1 Image

image details. It helps to compare image quality among clean, noisy and denoised CT images. The line segment of CT1 image is shown in Fig. 13, and intensity profile of clean, noisy, and denoised CT images are shown in Fig. 14. The intensity profile shows the clean image as smooth profile resulting better noise reduction without losing fine image details, and noisy image indicated as rapid, uneven fluctuations in intensity profile. Denoised (restored) images look more familiar to the clean

image profile, and in some situations, slightly over-smoothing, and loss of fine image features. The intensity profile of ResNet50-based CT image denoising model smooth out noise effectively, slightly suppresses sharp peaks, minimizing fine image detail. However, Over-smoothing in noisy areas can generate loss of fine image structures

and lowered visibility of small details but often clinically crucial features.

The intensity profile of UNet model retains overall CT image structures but smoothening rapid-intensity changes, reducing fine edge clarity, and skip connections are mainly used to preserve image details while suppressing noise.

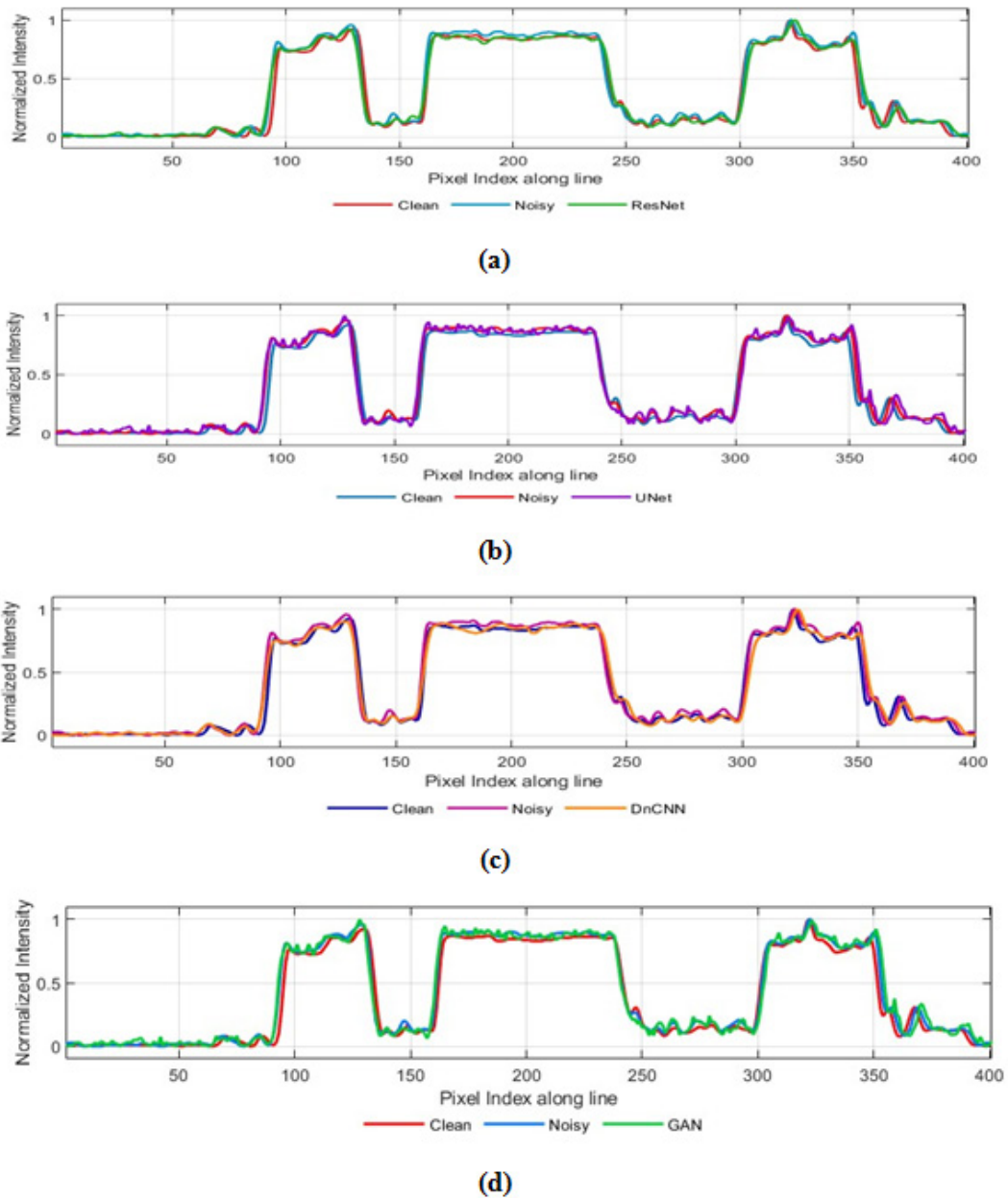


Fig. 14. Intensity profile of clean, noisy and various denoising methods

However, this method might fail to high-frequency changes, indicating as a softer appearance in in-depth anatomical regions. The intensity profile of DnCNN denoising model slightly follows the clean image intensity profile. It performs better in general noise reduction, but often varies at higher gradient regions, causing minute loss of fine details. The intensity profile of GAN-based model shows that the denoised image must closely matches the original clean image among other standard denoising methods. GAN model preserves the shaper peak (edge) details and deep valleys (structure) details without over-smoothing, resulting stronger preservation of structural information. Overall, the GAN model effectively reduces noise while preserving subtle variation in image tissue density, which is significant for medical interpretation.

Future work

Future enhancements of GAN-based CT image denoising, including real time application, and integration with diagnostic imaging tools to improve accurate diagnostic precision and expansion to other medical imaging techniques like (X-Rays, MRI, and PET). In the future, in terms of architecture, improvements such as integrating GAN with different models, such as diffusion or 3D architectures to improve computational efficiency in practical applications. Clinical integration, and compliance approval can promote this model to enhance patient diagnostic outcomes in practical real-world scenarios. Researchers are increasingly motivated to adopt composite, dynamically weighted loss functions that integrate perceptual, adversarial, and physics-informed features. Comparative studies of diverse GAN-based models and hybrid-domain training methods can further improve noise reduction while preserving fine anatomical structures. Validation should extend beyond metrics such as PSNR and SSIM, incorporating advanced perceptual metrics and expert healthcare assessments. Moreover, benchmarking both conventional and modern non-GAN denoising methods is essential to provide a comprehensive performance analysis and ensure improved diagnostic accuracy.

CONCLUSION

The denoising efficiency of the hybrid

GAN-based model was examined for lower dose CT images, and its outcomes were compared with those of the ResNet50³⁵, UNet³⁶, DnCNN³⁷ and GAN-based model for noise mitigation in clinical images. These experimental outcomes in terms of quantitative, qualitative, and computational time complexity analysis are compared with standard denoising methods. The GAN-based model maintained fine anatomical details, unlike the other existing models. Overall, the GAN-based model gives high fidelity and an optimal trade-off among noise mitigation and fine feature conservation. These experimental outcomes demonstrate an improvement in LDCT imaging quality over other existing models, indicating increased diagnostic precision and less radiation exposure in patients. The effective utilization of deep learning models and safer imaging approaches in medical images resulting better diagnostic outcomes.

ACKNOWLEDGEMENT

This research work is the final result of a collective effort among all the listed authors, each of whom contributed considerably to its preparation and completion.

Funding Sources

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest

The author(s) do not have any conflict of interest.

Data Availability

This statement does not apply to this article.

Ethics Statement

This research did not involve human participants, animal subjects, or any material that requires ethical approval.

Informed Consent Statement

This study did not involve human participants, and therefore, informed consent was not required.

Clinical Trial Registration

This research does not involve any clinical trials

Permission to reproduce material from other sources

Not applicable.

Author contributions

Swapna Katta: Conceptualization, Data Collection, Methodology, experimentation, Writing- original Draft; Prof. Deepak Garg: Conceptualization, Editing; Dr. Prabhishek Singh: Supervision, experimentation and validation; Prof. Manoj Diwakar: Visualization, Supervision.

REFERENCES

1. Diwakar M, Kumar M. A review on CT image noise and its denoising. *Biomed Signal Process Control*. 2018;42: 73–88.
2. Zhang C, Yen KS. A hybrid frequency–spatial domain unsupervised denoising model for Gaussian–Poisson mixed noise in medical imaging. *Computers in Biology and Medicine*. 2025;193:110374.
3. Diwakar M, Kumar M. CT image denoising using NLM and correlation based wavelet packet thresholding. *IET Image Process*. 2018;12(5):708–15.
4. Diwakar M, Singh P, Karetla GR et al. Low-dose COVID-19 CT image denoising using batch normalization and convolution neural network. *Electronics*. 2022;11(20):3375.
5. Zhang Z, Zhang P, Liu X, Hou J, Feng Q, Wang J. Image denoising via double-weighted correlated total variation regularization. *Applied Intelligence*. 2025;55(4):269.
6. Li S, Li Q, Li R, et al. An adaptive self-guided wavelet convolutional neural network with compound loss for low-dose CT denoising. *Biomedical Signal Processing and Control*. 2022;75: 103543.
7. Kaur R, Juneja M, Mandal AK. A comprehensive review of denoising techniques for abdominal CT images. *Multimedia Tools and Applications*. 2018;77: 22735–70.
8. Zeng GL. An attempt of directly filtering the sparse-view CT images by BM3D. *Proceedings of the 7th International Conference on Image Formation in X-Ray Computed Tomography*. 2022; 12304:568–75.
9. Toumoulin C. Improving abdomen tumor low-dose CT images using a fast dictionary learning-based processing.
10. Fan H, Lu M, Zhang X, et al. Wavelet-domain frequency-mixing transformer unfolding network for low-dose computed tomography image denoising. *Quantitative Imaging in Medicine and Surgery*. 2025;15(8):7419.
11. Katageri GS, Swamy PS. Denoising and analysis of synthetic aperture radar images using improved weight threshold technique in curvelet transform frequency domain. *Multimedia Tools and Applications*. 2025;84(12):10173-10194.
12. Sadia RT, Chen J, Zhang J. CT image denoising methods for image quality improvement and radiation dose reduction. *Journal of Applied Clinical Medical Physics*. 2024;25(2): e14270.
13. López-Aguirre M, El-Khatib-Núñez J. Neuroimaging applications in hepatic encephalopathy: acquisition protocols and basic preprocessing techniques. *In: Experimental and Clinical Methods in Hepatic Encephalopathy Research*. New York, NY: Springer US; 2025; 61–104.
14. Usui K, Ogawa K, Goto M, Sakano Y, Kyougoku S, Daida H. Quantitative evaluation of deep convolutional neural network-based image denoising for low-dose computed tomography. *Visual Computing for Industry, Biomedicine, and Art*. 2021;4:1–9.
15. Kumar RR, Priyadarshi R. Denoising and segmentation in medical image analysis: A comprehensive review on machine learning and deep learning approaches. *Multimedia Tools and Applications*. 2025;84(12):10817-10875.
16. Song H, Chen L, Cui Y, et al. Denoising of MR and CT images using cascaded multi-supervision convolutional neural networks with progressive training. *Neurocomputing*. 2022;469:354–365.
17. Ferdi A, Benierbah S, Nakib A. Residual encoder-decoder based architecture for medical image denoising. *Multimedia Tools and Applications*. 2024;1–18.
18. Yang Q, Yan P, Zhang Y, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE transactions on medical imaging*. 2018;37(6):1348–57.
19. Ali M, Ali M, Hussain M, Koundal D. Generative adversarial networks (GANs) for medical image processing: recent advancements. *Archives of Computational Methods in Engineering*. 2025;32(2):1185-1198.
20. Huang Z, Zhang J, Zhang Y, Shan H. DU-GAN: Generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. *IEEE Transactions on Instrumentation and Measurement*. 2021;71:1–12.
21. Wang G, Hu X. Low-dose CT denoising using a progressive Wasserstein generative adversarial network. *Computers in Biology and Medicine*. 2021; 135:104566.
22. Han Z, Shangguan H, Zhang X, Zhang P, Cui X, Ren H. A dual-encoder-single-decoder based low-dose CT denoising network. *IEEE Journal of Biomedical and Health Informatics*.

- 2022;26(7):3251–60.
23. Zhang X, Han Z, Shangguan H, Han X, Cui X, Wang A. Artifact and detail attention generative adversarial networks for low-dose CT denoising. *IEEE Transactions on Medical Imaging*. 2021;40(12):3901–18.
 24. Saidulu N, Muduli PR. Asymmetric convolution-based GAN framework for low-dose CT image denoising. *Computers in Biology and Medicine*. 2025;190: 109965.
 25. Kuraning V, Giraddi S, Baligar VP. Cycle-consistent generative adversarial network-based approach for denoising CT scan images. *Procedia Computer Science*. 2025;252:355–64.
 26. Abuya T, Rimiru R, Okeyo G. Adversarial content–noise complementary learning model for image denoising and tumor detection in low-quality medical images. *Signals*. 2025;6(2):17.
 27. Lokhande NL, Jaware TH. Effective CT lung image denoising using deep-dense inception generative adversarial network.
 28. He J, Deng J, Hu Z, Gu G, Qiao G. Structural semantic enhancement network for low-dose CT denoising. *Biomedical Signal Processing and Control*. 2025;108: 107870.
 29. Wang Y, Yang N, Li J. GAN-based architecture for low-dose computed tomography imaging denoising. *arXiv preprint arXiv:2411.09512*. 2024.
 30. Jiao F, Gui Z, Liu Y, Yao L, Zhang P. Low-dose CT image denoising via frequency division and encoder-dual decoder GAN. *Signal Image and Video Processing*. 2021;15(8):1907–15.
 31. Wang Y, Luo S, Ma L, Huang M. RCA-GAN: an improved image denoising algorithm based on generative adversarial networks. *Electronics*. 2023;12(22):4595.
 32. Naser MAU, Al-Asadi AHH. Image denoising using generative adversarial network by recursive residual group. *Journal of Robotics and Control (JRC)*. 2025;6(2):514-526.
 33. Wolterink JM, Leiner T, Viergever MA, Išgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Transactions on Medical Imaging*. 2017;36(12):2536–45.
 34. Maftouni M. Large COVID-19 CT slice dataset. *Kaggle*. Published 2021. <https://www.kaggle.com/datasets/maedemaftouni/large-covid19-ct-slice-dataset>.
 35. Shangguan T, Yang C, Zhang Y. ICU-Net: A U-shaped low-dose CT image denoising network based on codec structure. *Informatica*. 2025;49(6).
 36. Tripathi M, Kondo T. Optimizing image quality through UNET-based architectures in the denoising of facial and CT-scan images [dissertation]. *Thammasat University, Pathum Thani (Thailand)*. 2023.
 37. Singh P, Diwakar M, Gupta R, et al. A method noise-based convolutional neural network technique for CT image denoising. *Electronics*. 2022;11(21):3535.
 38. Hou H, Jin Q, Zhang G, Li Z. CT image quality enhancement via a dual-channel neural network with jointing denoising and super-resolution. *Neurocomputing*. 2022; 492:343-352.
 39. Lepcha DC, Goyal B, Dogra A, Vaghela K, Singh A, Kumar KR, Bavirisetti DP. Low-dose computed tomography image denoising using pixel level non-local self-similarity prior with non-local means for healthcare informatics. *Sci Rep*. 2025;15(1):25095.