ОРИГИНАЛЬНАЯ СТАТЬЯ

LOW-DOSE CT MEDICAL IMAGE DENOISING: NLM AND SIMPLE CNN PERFORMANCE COMPARISON

Ismail[1,2], Susilo Widodo[1], Brazovskiy K.S.[2]

1 – Research Center for Safety, Metrology, and Nuclear Quality Technology, Research Organization for Nuclear Energy, National Research and Innovation Agency. South Tangerang, Indonesia.

2 – National Research Tomsk Polytechnic University. Tomsk, Russia.

To reduce the noise level (denoising) of low-dose CT medical images using Non-Local Means (NLM) and Convolutional Neural Networks (CNN) and then compare the performance of both quantitatively and qualitatively.

Materials and Methods. A total of 856 image pairs (low-dose and full-dose) of DICOM images from patients with code L219, image size 512 × 512 pixels, and slice thickness of 1 mm, were downloaded from the American Association of Physicists in Medicine (AAPM) open dataset. The denoising methods used were NLM and CNN. The quantitative evaluation metrics used were Root Mean Square Error (RMSE), Peak-Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mahalanobis Distance. Qualitative evaluation is done by visual observation using IndoQCT software.

Results. The image quality resulting from the denoising process using CNN is better than NLM based on the results of calculating RMSE, PSNR, SSIM, and Mahalanobis Distance values.

Discussion. The denoising process using CNN provides better results. CNN is one of the deep learning models widely used for the denoising process. This model learns clear image patterns as ground truth and then applies the denoising process to noisy images. In carrying out the denoising process, patch-based NLM does not read the clear image pattern but replaces the pixel value with a weighted average of other pixels throughout the image itself. This weight is determined by the similarity of the patches so that detail and texture are still well maintained.

Conclusion. The CNN model successfully performed better on the denoising process, demonstrating the superiority of deep learning over classical denoising methods such as NLM. The simple CNN model proposed in this study also ran well on a standard-spec computer, opening up significant opportunities for researchers or students interested in pursuing this research topic.

Keywords: low-dose, computed tomography, medical image, denoising, NLM, CNN.

Corresponding author: Ismail, e-mail: ism22@tpu.ru, isma012@brin.go.id

For citation: Ismail, Susilo Widodo, Brazovskiy K.S. Low-dose CT medical image denoising: NLM and simple CNN performance comparison. REJR 2026; 16(1):158-164. DOI: 10.21569/2222-7415-2026-16-1-158-164.

Received: **18.10.25** **Accepted:** **12.11.25**

УСТРАНЕНИЕ ШУМА НА КТ-ИЗОБРАЖЕНИЯХ С НИЗКОЙ ДОЗОЙ ОБЛУЧЕНИЯ: СРАВНЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ НЕЛОКАЛЬНЫХ СРЕДНИХ (NLM) И СВЕРТОЧНЫХ (CNN) НЕЙРОННЫХ СЕТЕЙ

Исмаил[1,2], Сусило Видодо[1], Бразовский К.С.[2]

1 – Научно-исследовательский центр безопасности, метрологии и технологий качества ядерной энергии, Научно-исследовательская организация по атомной энергетике, Национальное агентство исследований и инноваций. Южный Тангеранг, Индонезия.
2 – Томский Национальный исследовательский политехнический университет. г. Томск, Россия.

Целью данного исследования является снижение уровня шума (шумоподавление) на медицинских КТ-изображениях с низкой дозой облучения с использованием нелокальных средних (NLM) и сверточных нейронных сетей (CNN), а затем сравнение производительности обоих методов в количественном и качественном плане.

Материалы и методы. Всего из открытого набора данных Американской ассоциации физиков-медиков (AAPM) было загружено 856 пар (с низкой и полной дозой) DICOM-изображений пациентов с кодом L219, размером изображения 512 × 512 пикселей и толщиной среза 1 мм. В качестве методов шумоподавления использовались нелокальные средние (NLM) и свёрточные нейронные сети (CNN). В качестве количественных метрик оценки использовались: среднеквадратическая ошибка (RMSE), отношение пикового сигнала к шуму (PSNR), индекс структурного сходства (SSIM) и расстояние Махаланобиса. Качественная оценка проводилась путем визуального наблюдения с использованием программного обеспечения IndoQCT.

Результаты. Качество изображения, полученное в результате шумоподавления с использованием CNN, лучше, чем в случае NLM, судя по результатам расчета значений RMSE, PSNR, SSIM и расстояния Махаланобиса.

Обсуждение. Процесс шумоподавления с использованием CNN обеспечивает лучшие результаты. CNN – одна из моделей глубокого обучения, широко используемых для шумоподавления. Эта модель распознает четкие паттерны изображения как истинные данные, а затем применяет процесс шумоподавления к зашумленным изображениям. При выполнении процесса шумоподавления NLM на основе патчей не считывает четкий паттерн изображения, а заменяет значение пикселя средневзвешенным значением других пикселей по всему изображению. Этот вес определяется сходством участков, что позволяет сохранить детализацию и текстуру.

Заключение. Модель CNN успешно справилась с шумоподавлением, продемонстрировав превосходство глубокого обучения над такими классическими методами шумоподавления, как NLM. Простая модель CNN, предложенная в данном исследовании, также хорошо работала на стандартном компьютере, открывая значительные возможности для исследователей и студентов, заинтересованных в изучении данной темы.

Ключевые слова: низкая доза облучения, компьютерная томография, медицинское изображение, шумоподавление, NLM, CNN.

Контактный автор: Исмаил, e-mail: ism22@tpu.ru, isma012@brin.go.id

Для цитирования: Исмаил, Сусило Видодо, Бразовский К.С. Устранение шума на КТ-изображениях с низкой дозой облучения: сравнение производительности нелокальных средних (NLM) и сверточных (CNN) нейронных сетей. REJR 2026; 16(1):158-164. DOI: 10.21569/2222-7415-2026-16-1-158-164.

Статья получена: **18.10.25** **Статья принята:** **12.11.25**

Research on low-dose CT medical image denoising and reconstruction using deep learning has emerged as a critical area of inquiry due to the increasing demand for reducing radiation exposure while maintaining diagnostic image quality [1]. Since the inception of computed tomography, advancements have evolved from traditional filtered back-projection and iterative reconstruction methods to sophisticated deep-learning based approaches. These developments are significant given that CT scans contribute substantially to medical radiation exposure, with low-dose CT (LDCT) techniques reducing radiation by up to 25% but introducing noise and artifacts that degrade image quality [2, 3]. The practical implications include lowering cancer risk and improving patient safety without compromising diagnostic accuracy.

The specific problem addressed is the challenge of effectively denoising LDCT images and reconstructing high-quality images despite the complex noise patterns and artifacts introduced by dose reduction. Existing methods, including sinogram filtering, iterative reconstruction, and image-domain post-processing, have limitations such as computational cost, proprietary data dependency, and loss of fine structural details [4]. Deep learning methods have shown promise but face knowledge gaps related to generalizability across different dose levels, interpretability, and the preservation of subtle anatomical features. Controversies exist regarding the balance between noise suppression and detail preservation, with some approaches leading to over-smoothing or artificial artefacts [5, 6]. The consequences of these gaps include potential misdiagnosis and limited clinical adoption of deep learning denoising techniques.

The conceptual framework integrates key concepts of low-dose CT imaging, noise and artifact characteristics, and deep learning-based denoising and reconstruction models. LDCT imaging reduces radiation dose but increase noise, necessitating advanced denoising algorithm. Deep learning model, including CNN, Generative Adversarial Network (GAN), and hybrid architectures, learn mappings from noisy to clean images, aiming to suppress noise while preserving structural integrity [7].

In this study, we aim to perform the denoising process using the NLM method as a conventional image processor and CNN as a deep-learning-based image processor, and compare the performance of both in performing the denoising process based on quantitative calculations through evaluation metrics: RMSE, PSNR, SSIM, and Mahalanobis Distance [8]. Meanwhile, qualitative image quality assessment is carried out by observation using IndoQCT [9].

Purpose.

This study aims to reduce the noise level (denoising) of low-dose CT medical images using Non-Local Means (NLM) and Convolutional Neural Networks (CNN) and then compare the performance of both quantitatively and qualitatively.

Materials and Methods.

A total of 856 image pairs (low dose and full dose) of DICOM image with a slice thickness of 1 mm and a size of 512 × 512 pixels were downloaded from the American Association of Physicists in Medicine (AAPM) open dataset [10]. The denoising methods used were NLM and CNN.

Mathematically, the NLM filter can be described aswhere the vector u ̂ represents the noisy image to be smoothed, k denoted the pixel index within the neighbouring window area of a target pixel for computational efficiency of pixel j, $w_{jk}(\hat{u})$ is the weighting coefficient and satisfies the condition $0 \le w_{jk}(\hat{u}) \le 1$ and $\sum_{k \in SW_j} w_{jk}(\hat{u}) = 1$, and $NLM(\hat{u}_j)$ denotes the intensity of pixel j after the NLM denoising.

In this case, $w_{jk}(\hat{u})$ calculated as

$$w_{jk}(\hat{u}) = \frac{\exp(-\|P(\hat{u}_j) - P(\hat{u}_k)\|^2_{2,a}/h^2}{\sum_{k \in SW_j} \exp(-\|P(\hat{u}_j) - P(\hat{u}_k)\|^2_{2,a}/h^2}$$

where $P(\hat{u}_j)$ and $P(\hat{u}_k)$ are patches around the pixel j and k, with $\|P(\hat{u}_j) - P(\hat{u}_k)\|$is the distance between two patches with a Gaussian kernel to weight the contribution to each dimension. Where h is the filtering parameter that controls the exponential decay and the weighting coefficient. When h is small, the image tends to be weakly smoothed; when h is large, the image tends to be strongly smoothed [11]. This research has applied two parameters in NLM denoising, h = 1.15 and 2.3 to control the smoothing level, patch size (ps) of 5 and 7, and patch distance (pd) of 6 and 15.

The CNN algorithm consists of an input layer, an output layer and several hidden layers of neurons, which perform an alternating sequence of linear and non-linear transformations. The output of each hidden layer is called a feature map. Each convolution layer is multiplied by the output of the previous layer. The activation function further processes the input multiplied by the weights of the convolution layer. The l-th input is given by

$$X_{ij} = f\left(\sum_{i \in M_j} x_i^{l-1} * w_{ij}^l + b_j^l\right)$$

Fig. 1 a (Рис. 1 а)

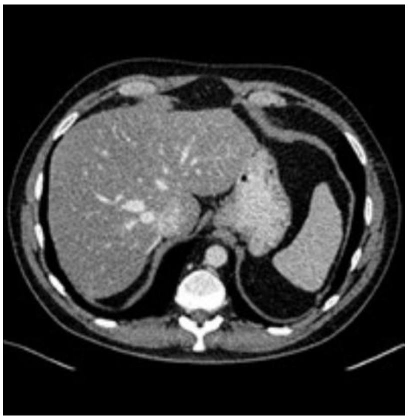

Fig. 1 b (Рис. 1 б)

Fig. 1 c (Рис. 1 в)

Fig. 1 d (Рис. 1 г)

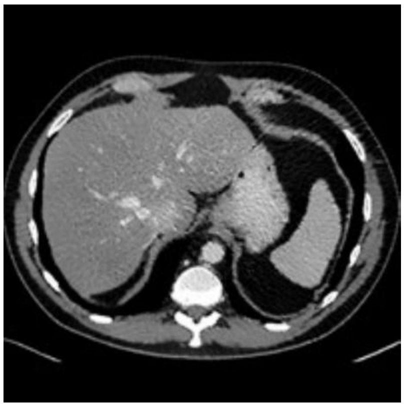

Fig. 1 e (Рис. 1 д)

Fig. 1. CT scan image, abdomibal area, number 315 observed using IndoQCT software.

a – Low-dose irradiation result,

b – Full-dose irradiation result,

c – Denoising result using simple CNN,

d – Denoising result using NLM A, h = 1.15, patch size = 5, patch distance = 6,

e – Denoising result using NLM B, h = 2.30, patch size = 7, patch distance = 15.

Рис. 1. КТ, органы брюшной полости, изображение номер 315, полученное с помощью программного обеспечения IndoQCT.

а – Результат облучения низкой дозой,

b – Результат облучения полной дозой,

c – Результат шумоподавления с использованием простой сверточной нейронной сети,

d – Результат шумоподавления с использованием NLM: h = 1,15, размер области = 5, расстояние между областями = 6,

e – Результат шумоподавления с использованием NLM: h = 2,3, размер области = 7, расстояние между областями = 15.

with M_j represents the selection of the input feature map, x_i^(l-1) is the output of the previous feature map, w_ij^l is the weight of the convolutional kernel of the l-th layer, f is the activation function which in this research used the leaky rectified linear unit (LeakyReLU), and b_j^l is the bias of the l-th layer [12]. In this research, simple CNN algorithm is built with an encoder section of two convolution layers, 16 and 32 filters, each followed by a LeakyReLU activation function, a bottleneck section that acts as a narrowing of the representation to extract the main features of the image, with Conv2D measuring 64 filters, and a decoder section that uses two Conv2DTranspose deconvolution layers with 32 and 16 filters to restore the image size. The model is trained using training data for 5 epochs and a batch size of 1, to adapting the computing capabilities of the available computer devices. We split the 80% dataset as training data and 20% as testing data, a random state of 42, using Adaptive Moment Estimation (Adam) optimizer, and Mean Square Error (MSE) as loss function. To build this model, we used Python 3.12 and TensorFlow 2.17 as the framework.

The quantitative evaluation metrics used were RMSE, PSNR, SSIM, and Mahalanobis Distance. Qualitative evaluation is done by visual observation using IndoQCT software. RMSE is a common image quality evaluation metric, used to measure the degree of difference between the output image and the original image. The smaller RMSE value indicates the smaller difference between the output image and the original image, showing the better the processing effect. PSNR is an objective measure of the error between image pixels and is typically used for error-sensitive images. The higher the PSNR value, the smaller the distortion and the better the image effect. Structural Similarity Index (SSIM) stands for structural similarity. It is an index used to measure the similarity between the full-dose and the denoised image, which lies in the range 0 to 1. It is better in line with human visual perception. A higher SSIM value close to 1 indicates better reconstruction quality, the more similar to the original [13].

This research also used quantitative analysis using Mahalanobis Distance. This distance indicates the difference between the two images being analysed. the smaller the distance value indicates that the two images are more similar [14].

Results and discussion.

A representative visual comparison of CT images for low dose, full dose, simple CNN denoised, and NLM denoised are shown in Figure 1. Visually, differences in noise levels were observed in the images. The low-dose irradiation images showed higher noise levels than the full-dose irradiation images. The denoising results using NLM and CNN also showed differences in noise levels. Compared to the low-dose irradiation images, a visual decrease in noise levels was observed, indicating successful denoising.

The NLM denoising images showed that the parameter values used significantly influenced the denoised images. Further quantitative analysis will be conducted using evaluation metrics such as RMSE, PSNR, SSIM, and Mahalanobis Distance. The results of the average evaluation metrics calculation of 856 images are shown in Table 1.

Table №1. The results of the average evaluation metrics calculation of 856 images.

RMSE				PSNR			
Noisy	NLM A	NLM B	CNN	Noisy	NLM A	NLM B	CNN
40.757	38.398	29.288	26.412	35.676	36.204	38.551	39.381

SSIM				Mahalanobis Distance			
Noisy	NLM A	NLM B	CNN	Noisy	NLM A	NLM B	CNN
0.845	0.868	0.917	0.931	247871.414	214600.146	130803.035	103544.510

NLM A = using parameters h = 1.15, patch size = 5, patch distance = 6

NLM B = using parameters h = 2.30, patch size = 7, patch distance = 15

The results of the evaluation metrics calculation indicate a decrease in the noise level in the denoised image. The parameter settings in the denoising method using NLM significantly affect the noise level. Larger parameters such as h, patch size, and patch distance result in better denoising. This is indicated by a smaller RMSE value, an increased PSNR value, an SSIM value closer to 1, and a smaller Mahalanobis Distance value. This supports the results of visual observations. But, tuning parameter values in NLM requires trial and error. Setting a larger h value improves smoothing but risks losing detail, while larger patch sizes and search windows improve robustness to noise but increase computational costs. This makes it impractical to perform the denoising process automatically.

However, the best results from image denoising using NLM are not better than those from CNN. This is because the image denoising process of NLM and CNN is different. NLM performs patch-based denoising that replaces pixel values with a weighted average of other pixels across the image, where the weights are determined by the similarity of the patches so that detail and texture are better preserved. NLM takes advantage of pattern repetition in the image: similar patches may appear far from the target location, so averaging the values from similar patches will suppress random noise while preserving the actual structure.

CNN-based image denoising, on the other hand, learns an end-to-end mapping from a noisy image to a clean one using convolutional layers, often in a residual framework, to predict noise and then subtract it from the input. This approach preserves structural details better than classical techniques when trained with sufficient data, although it requires high computational effort and careful training setup. With low-dose-full-dose image pairs available, CNNs are able to predict better by learning patterns from the clean (full-dose) image as ground truth. Therefore, CNNs are able to provide better denoising results than classical denoising methods, despite using a simpler architecture. Table 1 clearly shows the best denoising results using CNN. The RMSE value is the smallest, indicating a small difference between the denoised image and the full-dose image. An increase in the PSNR value indicates small image distortion. The SSIM value is closer to 1, indicating similarity to the ground truth image. Finally, the Mahalanobis Distance value is the smallest, indicating the closeness between the denoised image and the clean image.

Conclusions.

The CNN model successfully performed better on the denoising process, demonstrating the superiority of deep learning over classical denoising methods such as NLM. This is in accordance with the results of the evaluation metrics calculation, the image denoised using the simple CNN model shows the smallest RMSE value, the highest increase in PSNR value, the SSIM value closest to 1, and the smallest Mahalanobis Distance value. The simple CNN model proposed in this study also ran well on a standard-spec computer, opening up significant opportunities for researchers or students interested in pursuing this research topic.

Conflicts of interest.

The authors declare no conflict of interest.

Funding.

This research received no external funding.

References:

1. Immonen E. et al. The use of deep learning towards dose optimization in low-dose computed tomography: A scoping review. 2022, W.B. Saunders Ltd. doi: 10.1016/j.radi.2021.07.010.

2. Jo G., Ahn C., Hong J.H. et al. 75% radiation dose reduction using deep learning reconstruction on low-dose chest CT. BMC Med Imaging. 2023; 23: 121 https://doi.org/10.1186/s12880-023-01081-8.

3. Diwakar M. et al. Low-Dose COVID-19 CT Image Denoising Using Batch Normalization and Convolution Neural Network. Electronics (Basel). 2022; 11 (3375): 2-17 doi: 10.3390/electronics11203375.

4. Nagata, M., Ichikawa, Y., Domae, K. et al. Application of Deep Learning-Based Denoising Technique for Radiation Dose Reduction in Dynamic Abdominal CT: Comparison with Standard-Dose CT Using Hybrid Iterative Reconstruction Method. J Digit Imaging. 2023; 36: 1578-1587. https://doi.org/10.1007/s10278-023-00808-x

5. Wang H., L. L. Li, J. Shang, J. Song, and B. Liu. Application of deep learning image reconstruction in low-dose chest CT scan. British Journal of Radiology. 2022, 95: 1133 doi: 10.1259/bjr.20210380.

6. Kim J. H., H. J. Yoon, E. Lee, I. Kim, Y. K. Cha, and S. H. Bak. Validation of deep-learning image reconstruction for low-dose chest computed tomography scan: Emphasis on image quality and noise. Korean J Radiol. 2021; 22 (1): 131-138 doi: 10.3348/kjr.2020.0116.

7. Sadia R.T., J. Chen, and J. Zhang. CT image denoising methods for image quality improvement and radiation dose reduction. 2024, John Wiley and Sons Ltd. doi: 10.1002/acm2.14270.

8. Sara U., M. Akter, and M. S. Uddin. Image Quality Assessment through FSIM, SSIM, MSE and PSNR. A Comparative Study. Journal of Computer and Communications. 2019; 07 (03): 8-18 doi: 10.4236/jcc.2019.73002.

9. Anam C. et al. Assessment of patient dose and noise level of clinical CT images: Automated measurements. Journal of Radiological Protection. 2019; 39 (3): 783-793 doi: 10.1088/1361-6498/ab23cc.

10. Wang F., R. Wang, and H. Qiu. Low-dose CT reconstruction using datasetfree learning. PLoS One. 2024; 19 (6): 10.1371/journal.pone.0304738.

11. Zhang H., D. Zeng, H. Zhang, J. Wang, Z. Liang, and J. Ma. Applications of nonlocal means algorithm in low-dose X-ray CT image processing and reconstruction: a review. Med Phys. 2017; 44 (3): 1168-1185 doi: 10.1002/MP.12097.

12. Thakur R. S., R. N. Yadav, and L. Gupta. State-of-art analysis of image denoising methods using convolutional neural networks. 2019; 14. Institution of Engineering and Technology. doi: 10.1049/iet-ipr.2019.0157.

13. Li Z., S. Zhou, J. Huang, L. Yu, and M. Jin. Investigation of Low-Dose CT Image Denoising Using Unpaired Deep Learning Methods. IEEE Trans Radiat Plasma Med Sci. 2021; 5 (2): 224-234 doi: 10.1109/TRPMS.2020.3007583.

14. César Pereira J. P., A. Monteiro, R. D. L. Ribeiro, A. C. Sobieranski, and A. Von Wangenheim. Comparison of Supervised Classifiers and Image Features for Crop Rows Segmentation on Aerial Images. Applied Artificial Intelligence. 2020; 34 (4): 271-291.