ISSN (Print) 1226-1750
ISSN (Online) 2233-6656
Journal of Magnetics 31(3), 376-384 (2026) https://doi.org/10.4283/JMAG.2026.31.3.376

Multi-Level Wavelet Decomposition for Deep Learning-Based Medical Image Restoration: Application to Low-Dose CT

Minseung Kwak[1] and Byungdu Jo[1,2,3]*

[1]*Department of Multidisciplinary Radiological Sciences, Dongseo University, Busan 47011, Republic of Korea*
[2]*Department of Radiological Sciences, Dongseo University, Busan 47011, Republic of Korea*
[3]*Center for Radiological Environment & Health Science, Dongseo University, Busan 47011, Republic of Korea*

(Received 14 August 2026, Received in final form 4 September 2026, Accepted 7 September 2026)

Medical image restoration is important for maintaining image quality under reduced signal-to-noise conditions in medical imaging, including magnetic resonance imaging (MRI), in which magnetism is fundamental to image formation. Wavelet-based restoration can separate image information into different frequency components; however, conventional single-level decomposition may limit the representation of noise and structural information across multiple spatial scales. Therefore, this study investigated the effect of wavelet decomposition depth on deep learning-based restoration using computed tomography (CT) images under simulated low-dose conditions. A single-level wavelet cascade U-Net (SL-WC-U-Net) and a multi-level wavelet embedded U-Net (ML-WE-U-Net) were compared using Gaussian noise levels from 0.01 to 0.10. Restoration performance was evaluated using peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), visual inspection, absolute difference maps, and frequency-domain analysis. The ML-WE-U-Net consistently achieved higher PSNR and SSIM values, with PSNR improvements of 4.08–6.88 dB, and better preserved structural and frequency information. These findings suggest that multi-level wavelet decomposition can improve the balance between noise suppression and preservation of structural and frequency information, providing a potential image-restoration framework applicable to medical imaging, including magnetics-related imaging modalities.

Keywords : low-dose CT, medical image restoration, discrete wavelet transform, deep learning, electromagnetic field radiation, magnetism in medical imaging

1. Introduction

Medical imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) are widely used to visualize internal anatomical structures. Although their image formation principles differ, image quality in both modalities can be affected by reduced signal-to-noise conditions. In MRI, accelerated acquisition may decrease the signal-to-noise ratio (SNR), whereas in CT, reducing the radiation dose generally increases image noise and may decrease image contrast. These limitations have motivated the development of image restoration methods that suppress noise while preserving diagnostically relevant structural information. In particular, low-dose CT (LDCT) has been widely investigated as a means of reducing radiation exposure, although the resulting increase in noise can degrade the visibility of fine anatomical structures [1, 2]. Therefore, effective image restoration is important for improving image quality under low-SNR or noise-degraded imaging conditions.

Conventional noise reduction methods have primarily relied on image filtering and iterative reconstruction techniques. More recently, deep learning-based approaches have shown promising performance in medical image restoration by learning mappings between noise-degraded and reference images [3–8]. Among these approaches, U-Net has been widely used because its encoder–decoder architecture extracts hierarchical features, while skip connections transfer spatial information from the encoder to the corresponding decoder stages [9–11]. However, conventional U-Net architectures mainly process image features in the spatial domain and do not explicitly separate information according to frequency characteristics. As a result, noise suppression may also attenuate fine structural details during the restoration process.

To address this limitation, wavelet transforms have

©The Korean Magnetics Society. All rights reserved.
*Corresponding author: Tel: +82-51-320-4274
E-mail: byungdujo@gdsu.dongseo.ac.kr

© 2026 Journal of Magnetics

been incorporated into deep learning-based restoration models. The discrete wavelet transform (DWT) decomposes an image into four subbands: a low-frequency approximation subband (LL) and three high-frequency detail subbands (LH, HL, and HH). The LL subband mainly contains coarse structural information, whereas the detail subbands contain directional high-frequency components associated with edges, fine structures, and noise. By separating image information across different frequency components, DWT provides a means of suppressing noise while retaining structural details [12].

Previous studies have demonstrated that incorporating wavelet processing into deep learning networks can improve image restoration performance, and that the way in which wavelet operations are combined with the network influences noise reduction characteristics [12, 13]. However, many previous approaches have mainly employed single-level wavelet decomposition, limiting frequency analysis to a single spatial scale. Because DWT can be recursively applied to the LL subband, image information can be represented at multiple spatial resolutions. Such multi-level decomposition may provide a more effective representation of noise and structural information distributed across different spatial scales. At the same time, excessive decomposition may attenuate fine details and increase image blurring [14]. Therefore, the appropriate depth of wavelet decomposition remains an important factor in wavelet-based medical image restoration.

Accordingly, this study investigated the effect of wavelet decomposition depth on deep learning-based medical image restoration by comparing a single-level wavelet cascade U-Net (SL-WC-U-Net) with a multi-level wavelet embedded U-Net (ML-WE-U-Net) using CT images under simulated low-dose conditions. The comparison was designed to determine whether multi-level decomposition provides advantages not only in numerical image quality but also in the preservation of spatial and frequency-domain information. Restoration performance was therefore evaluated using peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), visual assessment, absolute difference maps, and frequency-domain analysis.

2. Materials and Methods

2.1. Image Dataset

The image dataset used in this study was the same as that employed in our previous work to ensure consistency in the experimental conditions. Publicly available chest and abdominal CT images obtained from Kaggle were used [11, 12, 15, 16]. All images were provided in DICOM format and consisted of thoracic and abdominal CT scans. Chest and abdominal CT images were combined into a single dataset and used together to train each noise-level-specific model, without anatomical region-specific separation. Normal-dose CT images were designated as reference images, and simulated low-dose CT (LDCT) images were generated by adding artificial noise to the corresponding normal-dose images. All images used in this study were obtained from a publicly available, de-identified dataset. No patients were recruited, no identifiable personal information was accessed, and no new clinical data were collected. Therefore, institutional review board approval and informed consent were not required for this study.

2.2. Preprocessing and Noise Simulation

All CT images were resized to 256×256 pixels and independently normalized to the range [0, 1]. Simulated LDCT images were generated by adding Gaussian noise with a mean of 0.05 and standard deviations ranging from 0.01 to 0.10 in increments of 0.01. At each noise level, five independent noise realizations were generated and evaluated. The mean PSNR and SSIM values were used for performance evaluation to reduce variability caused by random noise generation.

2.3. Wavelet Transform

The wavelet transform is a multiresolution analysis technique that represents a signal in both spatial and frequency domains. In this study, the discrete wavelet transform (DWT) was used to decompose CT images into low- and high-frequency components. The wavelet coefficients of a signal $f(t)$ can be expressed as

$$a_{j,k} = \int f(t)\psi_{j,k}(t)dt \tag{1}$$

where $\psi_{j,k}(t)$ denotes the wavelet basis function, with j and k representing the scale and translation parameters, respectively. The original signal can be reconstructed using the inverse transform,

$$f(t) = \Sigma_{j,k}\, a_{j,k}\psi_{j,k}(t) \tag{2}$$

where $a_{j,k}$ denotes the corresponding wavelet coefficient [14, 17].

For two-dimensional images, DWT applies low-pass and high-pass filters in the horizontal and vertical directions, generating four subbands: LL, LH, HL, and HH. The LL subband contains low-frequency approximation information, whereas the LH, HL, and HH subbands contain high-frequency detail information associated with

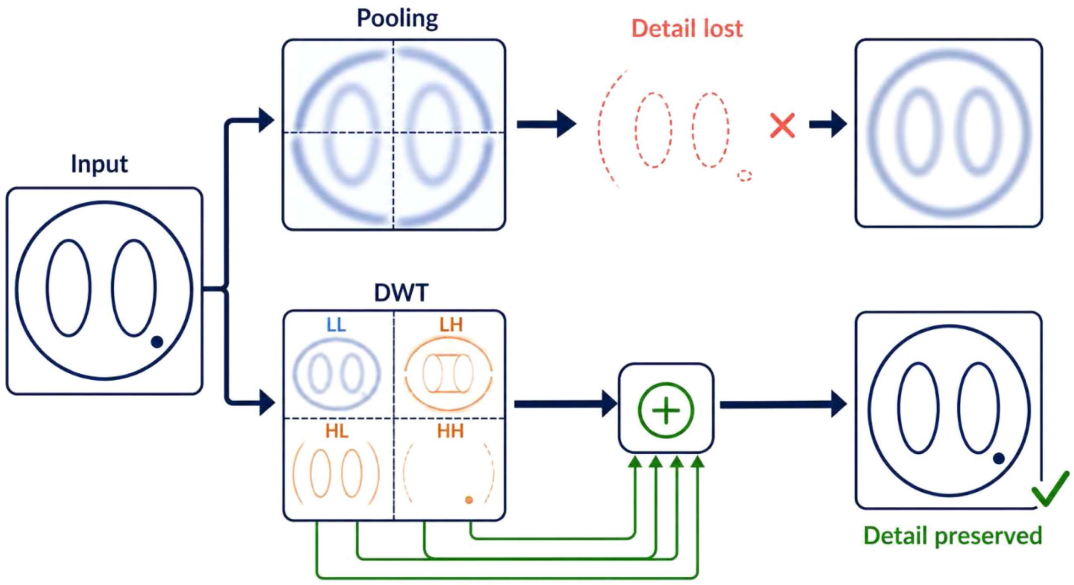


Fig. 1. (Color online) Conceptual comparison of pooling and DWT-based feature preservation.

edges, fine structures, and noise [18].

In the single-level configuration, DWT was applied once. In the multi-level configuration, the LL subband was recursively decomposed at successive encoder stages, while the corresponding detail subbands were retained for reconstruction in the decoder. This process enabled image information to be represented at multiple spatial resolutions.

The image signal can therefore be represented by low- and high-frequency components as

$$f(t) = \{c(t), d(t)\},\ \{c(t) = \text{low frequency},\ d(t) = \text{high frequency}\} \quad (3)$$

where $c(t)$ and $d(t)$ denote the low- and high-frequency coefficients, respectively. Fig. 1 illustrates the conceptual difference between conventional pooling and DWT-based decomposition [18]. The single-level and multi-level processing architectures used in this study are shown in Figs. 2 and 3, respectively.

2.4. U-Net

U-Net is an encoder–decoder convolutional neural network widely used for medical image restoration. The encoder progressively extracts hierarchical features while reducing spatial resolution, whereas the decoder restores

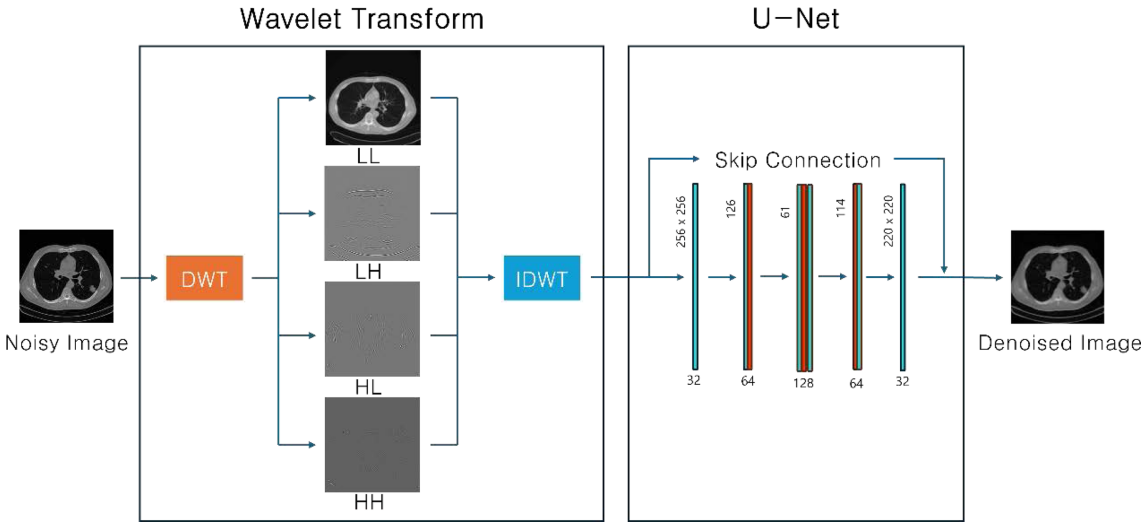


Fig. 2. (Color online) Architecture of the SL-WC-U-Net Model.

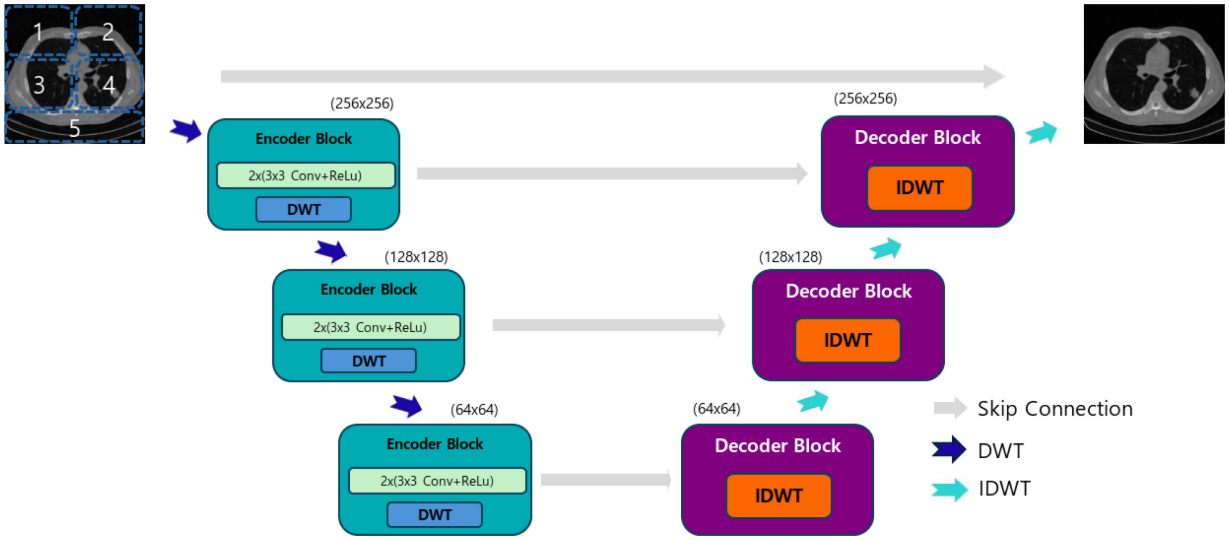


Fig. 3. (Color online) Architecture of the ML-WE-U-Net Model.

spatial resolution. Skip connections transfer feature maps from the encoder to the corresponding decoder stages, helping preserve spatial information during image restoration [8–10].

In the SL-WC-U-Net, a single-level DWT–IDWT operation was applied before U-Net, as shown in Fig. 2. In contrast, the ML-WE-U-Net recursively applied DWT across multiple encoder stages and restored spatial resolution through the corresponding IDWT stages in the decoder, as shown in Fig. 3. At each spatial scale, skip connections transferred feature information between the corresponding encoder and decoder blocks.

This design allowed the multi-level model to process wavelet information at multiple spatial resolutions, whereas the single-level model processed wavelet information at only one level. Both models were evaluated using the same dataset, preprocessing procedure, noise conditions, and evaluation metrics.

2.5. Training Configuration

Two experimental configurations were evaluated. The single-level configuration applied one DWT and IDWT operation before U-Net, whereas the multi-level configuration integrated three successive DWT and IDWT stages into the encoder–decoder pathway. Normalized noisy CT images were used as inputs, and the corresponding normalized reference images were used as targets. For each network architecture, the model was independently trained from scratch for each Gaussian noise level (σ = 0.01–0.10), resulting in a separate set of optimized weights for each noise level. For training the multi-level model, five 256 × 256-pixel patches were randomly extracted from each DICOM image, yielding a total of 500 patches. The training pairs consisted of noise-added input patches and their corresponding reference patches.

Training was performed using the Adam optimizer with a learning rate of 1×10^{-4}, a batch size of 8, and a maximum of 100 epochs. The data were divided into training and validation sets at a ratio of 8:2. Early stopping was applied based on the validation loss with a patience of 10 epochs, and the model with the lowest validation loss was retained. A weighted composite loss function was used to jointly optimize pixel-level accuracy, structural similarity, and edge preservation. The total loss was defined as $L_{total} = 0.4L_{MSE} + 0.5L_{SSIM} + 0.1L_{edge}$.

2.6. Performance Evaluation

PSNR and SSIM were calculated directly from the final reconstructed images generated by each network, without additional filtering, smoothing, thresholding, or other post-processing. Restoration performance was quantitatively evaluated using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). PSNR measures pixel-wise fidelity based on the mean squared error (MSE) between the reference and restored images, with higher values indicating lower reconstruction error.

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{4}$$

Because the images were normalized to the range [0, 1], the maximum pixel value used for PSNR calculation was 1. SSIM evaluates structural similarity by considering luminance, contrast, and structural information between the reference and restored images.

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (5)$$

SSIM ranges from 0 to 1, with values closer to 1 indicating greater structural similarity [19]. Together, PSNR and SSIM were used to assess pixel-level accuracy and structural preservation, respectively.

2.7. Difference Map Analysis

Absolute difference maps were calculated to evaluate the spatial distribution of residual errors after image restoration. For each image, the pixel-wise absolute difference from the reference image was calculated. Difference maps for the noisy, single-level, and multi-level images were displayed using the same intensity range of 0–0.12, allowing direct comparison of the location and magnitude of residual errors.

2.8. Frequency Spectrum Analysis

Two-dimensional fast Fourier transform (FFT) analysis was performed to evaluate the preservation of frequency components after image restoration. The zero-frequency component was shifted to the center of the spectrum, and the logarithmic magnitude spectrum was visualized. In addition, a radially averaged frequency profile was calculated from the centered spectrum to quantitatively compare the frequency distributions of the reference, noisy, single-level, and multi-level images. This analysis was used to assess whether the restoration models suppressed noise while preserving frequency information associated with image structures.

3. Results

PSNR and SSIM were evaluated for the SL-WC-U-Net and ML-WE-U-Net across Gaussian noise levels ranging from $\sigma = 0.01$ to 0.10. Five independent noise realizations were assessed at each noise level, and the mean values were used for comparison.

Table 1. PSNR and SSIM values for both models across noise levels.

Noise	SL-WC-U-Net		ML-WE-U-Net	
	PSNR	SSIM	PSNR	SSIM
0.01	35.2007	0.9449	42.0849	0.9765
0.02	34.5976	0.9302	39.3975	0.9625
0.03	33.6408	0.9123	38.8238	0.9523
0.04	32.9764	0.8867	38.4043	0.9450
0.05	32.6775	0.8699	37.4494	0.9327
0.06	31.7913	0.8578	36.5403	0.9234
0.07	31.5084	0.8294	36.0515	0.9117
0.08	31.3030	0.8302	35.8612	0.9067
0.09	30.6602	0.7951	34.7385	0.8967
0.1	30.5314	0.7962	35.0648	0.8869

Table 1 summarizes the PSNR and SSIM results for both models. The ML-WE-U-Net achieved higher PSNR and SSIM values than the SL-WC-U-Net at all noise levels. The PSNR of the ML-WE-U-Net decreased from 42.0849 dB at $\sigma = 0.01$ to 35.0648 dB at $\sigma = 0.10$, whereas that of the SL-WC-U-Net decreased from 35.2007 to 30.5314 dB. The corresponding PSNR improvement of the multi-level model ranged from 4.0783 to 6.8842 dB. A similar trend was observed for SSIM, with the ML-WE-U-Net consistently maintaining higher structural similarity across the evaluated noise levels.

As shown in Fig. 4, both PSNR and SSIM generally decreased as the noise level increased. Despite this decline, the ML-WE-U-Net maintained higher values than the SL-

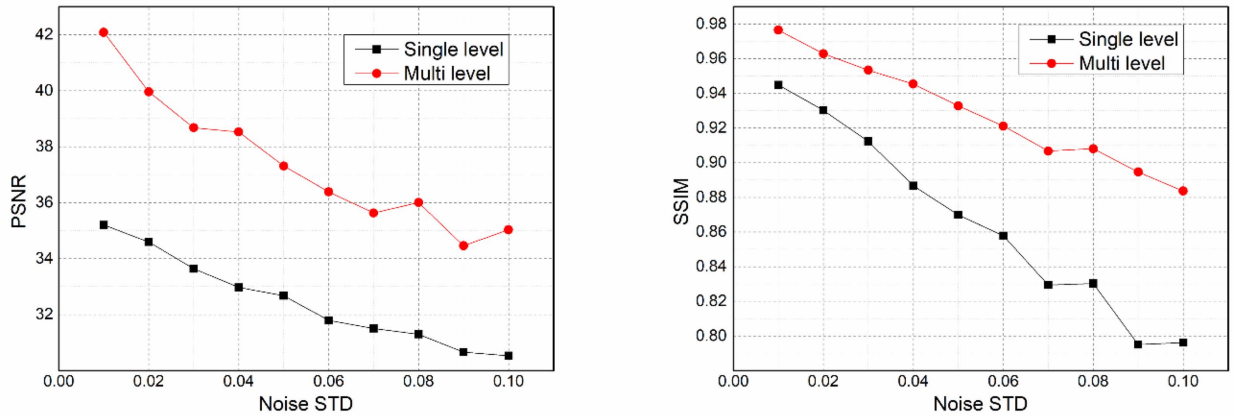


Fig. 4. (Color online) PSNR and SSIM trends for both models across noise levels.

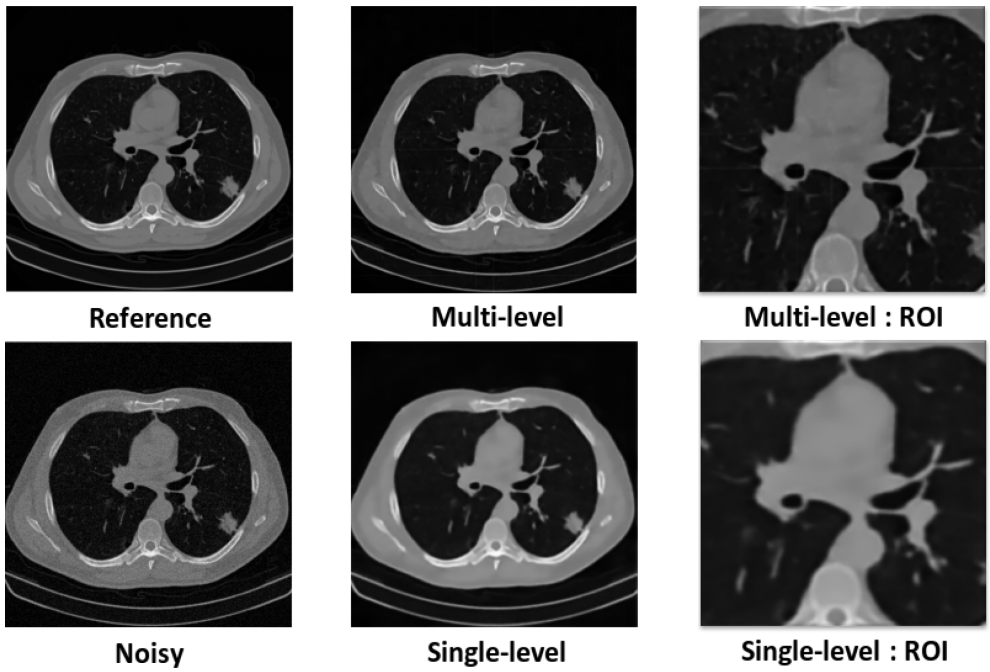


Fig. 5. Visual comparison of restored images and enlarged ROI (σ = 0.05).

WC-U-Net throughout the entire noise range, indicating more stable restoration performance under increasing noise conditions.

Figure 5 presents the reference image, noisy image, restored images, and enlarged regions of interest at σ = 0.05. Compared with the SL-WC-U-Net, the ML-WE-U-Net more effectively suppressed visible noise while retaining sharper anatomical boundaries and local structural variations. In contrast, the SL-WC-U-Net showed greater smoothing in the enlarged region. Although some localized differences from the reference image remained near high-contrast boundaries, the visual findings were consistent with the PSNR and SSIM results.

To further examine the spatial distribution of residual errors, absolute difference maps are shown in Fig. 6. The noisy image exhibited broadly distributed errors throughout the image. After restoration, the residual errors became more concentrated around structural boundaries. The ML-WE-U-Net showed lower residual differences in relatively homogeneous regions than the SL-WC-U-Net, although localized errors remained along high-contrast edges.

Frequency-domain characteristics were further evaluated

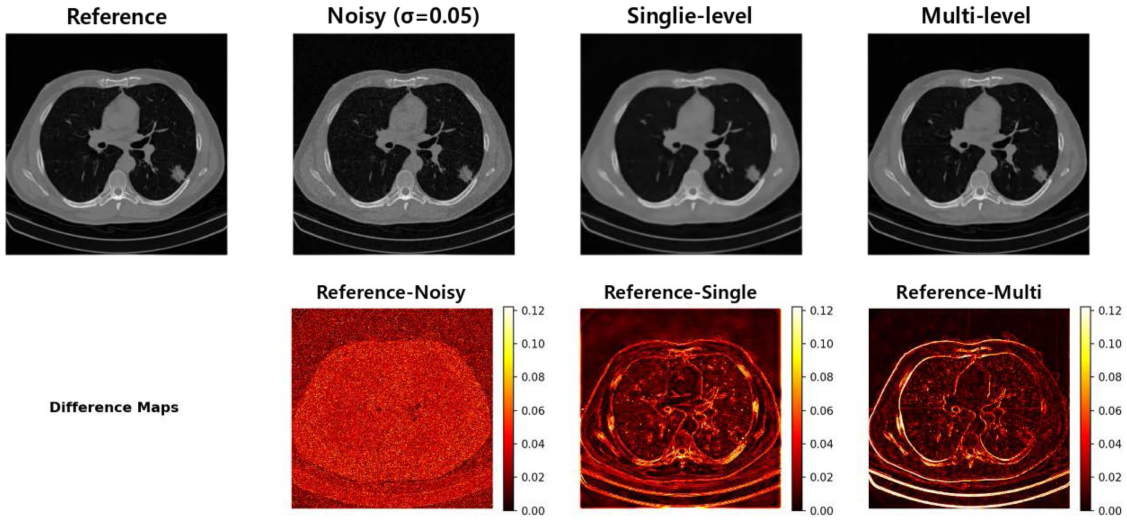


Fig. 6. (Color online) Absolute difference maps relative to the reference image.

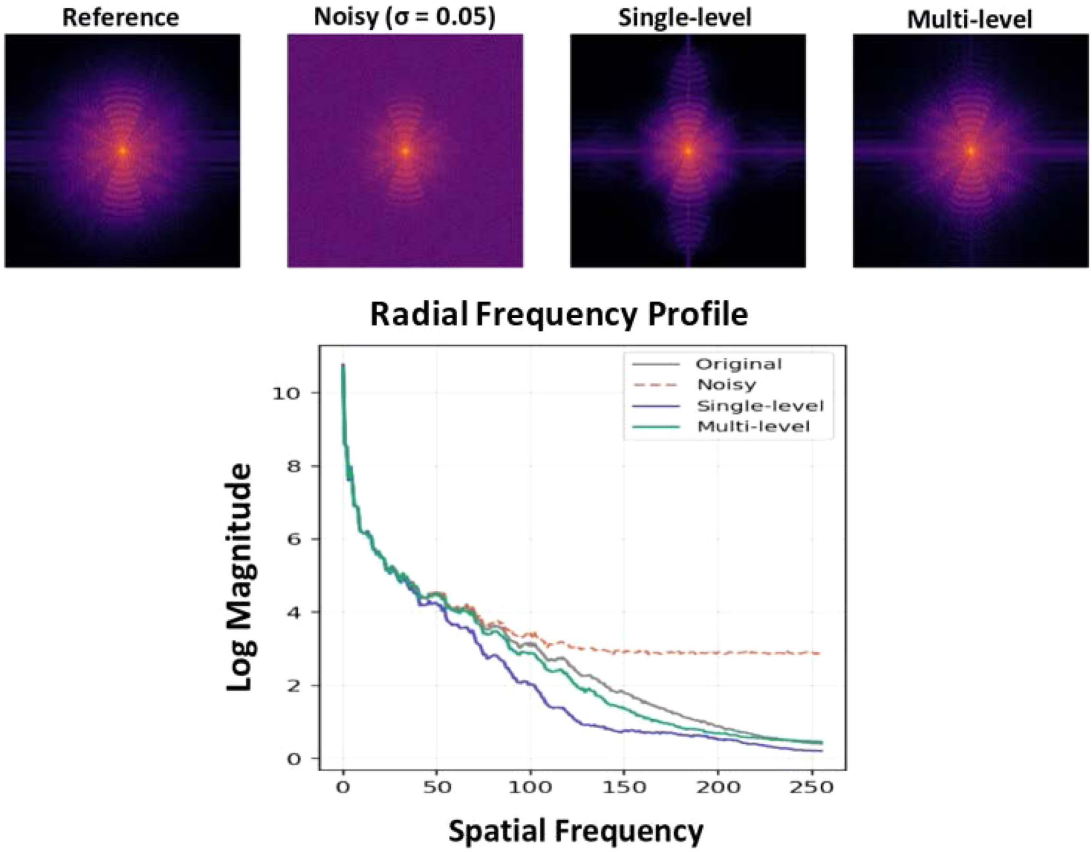


Fig. 7. (Color online) Frequency spectra and radial frequency profiles of the compared images.

using the two-dimensional Fourier magnitude spectrum and radially averaged frequency profiles, as shown in Fig. 7. The noisy image exhibited an elevated magnitude over the high-frequency range, reflecting increased noise components. The SL-WC-U-Net attenuated the mid- and high-frequency components below the reference profile, suggesting greater loss of high-frequency information. In contrast, the ML-WE-U-Net produced a radial frequency profile that more closely followed the reference image over a broader frequency range.

Overall, the quantitative, spatial, and frequency-domain analyses showed a consistent tendency. Compared with the single-level model, the multi-level model provided stronger noise suppression while maintaining structural and frequency information more effectively. However, residual differences and slight blurring remained in some fine structural regions, indicating that improved noise suppression does not necessarily ensure complete preservation of fine structural details [20, 21].

4. Discussion

In this study, the SL-WC-U-Net and ML-WE-U-Net models were compared to investigate the effect of multi-level wavelet processing on LDCT image restoration. The multi-level model consistently achieved higher PSNR and SSIM values across all evaluated noise levels and showed better preservation of structural information in the visual comparison.

The improved restoration performance may be attributed to the hierarchical frequency representation generated by repeated DWT and IDWT operations. In the single-level model, wavelet information is processed at only one spatial scale before U-Net-based restoration. In contrast, the multi-level model integrates low- and high-frequency information across multiple spatial resolutions. This multi-scale representation may enable the network to better distinguish noise-related components from structural features represented across different spatial scales.

The visual comparison and absolute difference maps were consistent with the quantitative results. The noisy images exhibited diffuse errors throughout the image, whereas the residual errors after restoration were mainly concentrated around anatomical boundaries. Compared with the single-level model, the multi-level model reduced residual errors more effectively in relatively homogene-

ous regions, although localized discrepancies remained along high-contrast edges. These findings suggest that multi-level wavelet processing improves overall noise suppression but does not completely eliminate errors around fine structural boundaries.

The frequency-domain analysis further supported these observations. The noisy image showed elevated high-frequency magnitudes, reflecting increased noise components. In contrast, the single-level output showed attenuation below the reference profile in the mid- and high-frequency ranges, indicating a greater loss of frequency information. The multi-level output remained closer to the reference profile over a broader frequency range, suggesting a better balance between noise suppression and preservation of structural frequency components.

Taken together, these results indicate that multi-level wavelet processing influences not only pixel-wise restoration accuracy but also the spatial distribution of residual errors and the preservation of frequency information. Therefore, PSNR and SSIM alone may not fully characterize restoration performance, and their interpretation should be complemented by spatial error analysis and frequency-domain evaluation when assessing LDCT image restoration models.

This study has several limitations. First, the limited CT dataset and the use of simulated Gaussian noise may not fully represent the complex noise characteristics of clinical LDCT images. Second, the training and validation images were obtained from the same dataset, and no independent external test dataset was included. Third, the input-sampling strategies differed between the two models: the SL-WC-U-Net was trained using full images, whereas the ML-WE-U-Net was trained using 500 patches extracted from 100 CT images. This difference may have partially influenced the computational cost and restoration performance in addition to the difference in wavelet decomposition depth. Fourth, only the db4 wavelet and two decomposition configurations were evaluated, limiting the assessment of other wavelet bases and decomposition depths. Finally, no subjective image quality assessment by radiologic technologists or radiologists was performed. Future studies should therefore evaluate additional wavelet bases and decomposition depths, incorporate more realistic LDCT noise conditions and patient-level independent datasets, report variability and statistical significance, and include task-based or expert image quality assessments.

5. Conclusion

This study investigated the effect of wavelet decomposition depth on deep learning-based medical image restoration by comparing the SL-WC-U-Net and ML-WE-U-Net under simulated low-dose CT conditions. The multi-level model consistently achieved higher PSNR and SSIM values across all evaluated noise levels, with PSNR improvements ranging from 4.08 to 6.88 dB compared with the single-level model.

Visual comparisons and absolute difference maps showed that the multi-level model more effectively suppressed diffuse noise while preserving structural information. In addition, frequency-domain analysis demonstrated that the radial frequency profile of the multi-level model remained closer to that of the reference image, indicating improved preservation of frequency components.

Overall, these findings suggest that multi-level wavelet decomposition can improve the balance between noise suppression and preservation of structural and frequency information in deep learning-based medical image restoration. However, localized errors remained near fine structural boundaries. Further validation using independent clinical LDCT datasets, additional wavelet bases and decomposition depths, statistical analysis, and expert image quality assessment is required to establish the robustness and clinical applicability of this approach.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2026-25596654).

References

[1] R. Rashma, J. Kumar, A. Garg, R. Batra, R. Meher, and A. Phulia, Egypt. J. Radiol. Nucl. Med. **55**, 85 (2024).
[2] J. Yuan, F. Zhou, Z. Guo, X. Li, and H. Yu, J. Digit. Imaging **36**, 2290 (2023).
[3] R. Heckel, M. Jacob, A. Chaudhari, O. Perlman, and E. Shimron, Magn. Reson. Mater. Phys. Biol. Med. **37**, 335 (2024).
[4] S. Kiryu, H. Akai, K. Yasaka, T. Tajima, A. Kunimatsu, N. Yoshioka, M. Akahane, O. Abe, and K. Ohtomo, Radiographics **43**, e220133 (2023).
[5] L. R. Koetzier, D. Mastrodicasa, T. P. Szczykutowicz, N. R. van der Werf, A. S. Wang, V. Sandfort, A. J. van der Molen, D. Fleischmann, and M. J. Willemink, Radiology **306**, e221257 (2023).
[6] J. Greffier, S. Si-Mohamed, J. Frandon, M. Loisy, F. de Oliveira, J. P. Beregi, and D. Dabli, Med. Phys. **49**, 5052 (2022).
[7] A. C. David-Olawade, D. B. Olawade, L. Vanderbloemen, O. B. Rotifa, S. C. Fidelis, E. Egbon, A. O. Akpan,

S. Adeleke, A. Ghose, and S. Boussios, J. Diagn. **15**, 689 (2025).
[8] O. Ronneberger, P. Fischer, and T. Brox, Lect. Notes Comput. Sci. **9351**, 234 (2015).
[9] H. Tang, N. Que, Y. Tian, M. Li, A. Perelli, and Y. Teng, Phys. Med. Biol. **70**, 045021 (2025).
[10] Y. Gu, ITM Web Conf. **73**, 02019 (2025).
[11] C. Y. Park and B. D. Jo, J. Magnetics **30**, 802 (2025).
[12] M. S. Kwak, C. Y. Park, and B. D. Jo, J. Korean Magn. **35**, 324 (2025).
[13] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, IEEE Access **7**, 74973 (2018).
[14] S. G. Mallat, IEEE Trans. Pattern Anal. Mach. Intell. **11**, 674 (1989).
[15] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, and F. Prior, J. Digit. Imaging **26**, 1045 (2013).
[16] B. Albertina, M. Watson, C. Holback, R. Jarosz, S. Kirk, Y. Lee, K. Rieger-Christ, and J. Lemmerman, The Cancer Genome Atlas Lung Adenocarcinoma Collection (TCGA-LUAD) (2016).
[17] M. J. Shensa, IEEE Trans. Signal Process **40**, 2464 (1992).
[18] M. A. Gungor, Optik **235**, 166652 (2021).
[19] M. Arabboev, S. Begmatov, M. Rikhsivoev, K. Nosirov, and S. Saydiakbarov, Acta IMEKO **13**, 1 (2024).
[20] G. Luo, D. Osypiw, and C. Hudson, Proc. SPIE. **6063**, 606303 (2006).
[21] S. G. Chang, B. Yu, and M. Vetterli, IEEE Trans. Image Process **9**, 1532 (2000).