# An Enhanced Noise Removal-based SAR Image Recognition using DnCNN and Wavelet Transform

*Youngdoo CHOI [1], Geunhwan KIM [2], Bong-seok KIM [3], Sangdong KIM [3]*

[1] Dept. of Electronics and Control Engineering, ROK Navy Academy, Jungwon-ro, Changwon-si, Republic of Korea
[2] Dept. of Electronic Engineering, Changwon National University (CWNU), Changwon, Republic of Korea
[3] Division of Automotive Technology, DGIST, 333 Techno Jungang-daero, Daegu, Republic of Korea

chododo78@navy.ac.kr, kimgw200@changwon.ac.kr, {remnant, kimsd728}@dgist.ac.kr

Submitted October 31, 2024 / Accepted March 27, 2025 / Online first June 16, 2025

**Abstract.** *This paper presents an enhanced method for noise removal and target detection in Synthetic Aperture Radar (SAR) images using a Denoising Convolutional Neural Network (DnCNN) combined with wavelet transform. Unlike conventional method, the proposed framework focuses on remove the Speckle Noise through residual learning and wavelet transform. The DnCNN architecture, consisting of 29 layers, efficiently removes noise while preserving high-frequency image features. The integration of wavelet transform further enhances noise removal and feature preservation. Experimental results demonstrate that the recognition rate of the proposed method improves by about 94% compared to original method under 10 dB Speckle Noise conditions. This method outperforms conventional algorithm in SAR image processing, making it highly suitable for applications in noisy environments.*

## Keywords

Navy SAR, noise, Convolutional Neural Network (CNN), Denoising Convolutional Neural Network (DnCNN), wavelet transform

## 1. Introduction

The Navy is using various technologies to monitor the marine environment and maintain marine safety. Among them, the Synthetic Aperture Radar (SAR) technology is being used as an important tool for detecting and analyzing marine conditions in real time among them [1]. The SAR radar can monitor the marine environment, coastal boundaries, naval facilities, and activities. It can also monitor maritime traffic and detect maritime transactions or pirate activities and can be used in lifesaving operations in the event of a disaster or accident. In military operations, the SAR radar also plays a key role. It monitors the location and movement of the enemy and detects the activities of the enemy, which helps in strategic decision-making. Unlike other sensors, the SAR radar has the advantage to use it even in dark conditions or severe weather, and thus increases high demand in all military including the Navy.

The operating principle of the SAR radar operates by transmitting electromagnetic waves to receive the signal reflected from the target of interest, and to reconstruct the image of the target. SAR radar systems are generally widely used as they are mounted on moving platforms such as aircraft or spacecraft [1]. The generation of noises such as Gaussian noise and speckle noise due to various reasons such as maximum distance makes it difficult to obtain clean images from SAR radars. In particular, noise is generated by system noise of radar and is recognized as a major problem in image radar systems [2]. In particular, Gaussian noise in SAR image radar images was removed through digital filters or wavelet conversion because distribution was difficult to predict [3], [4]. However, conventional filters are difficult to remove complex noise patterns, and noise cannot be effectively removed.

Recently, deep learning-based SAR image recognition has attracted attention, and in particular, improvements in noise cancellation techniques are playing an important role in improving recognition performance [5]. Effective noise cancellation is crucial, as SAR images inherently contain significant noise. Among the conventional deep learning-based noise cancellation techniques, noise cancellation convolutional neural networks (DnCNN) [6] and Pix2Pix [7], [8] have been widely used. However, these techniques show limitations in effectively removing complex noise patterns while maintaining structural details of SAR images.

To address this, this paper focuses on a noise cancellation technique that combines DnCNN and wavelet transform. By combining the strengths of the two methods, the technique achieves superior noise cancellation performance compared to conventional techniques. Experimental results show that the proposed technique outperforms the state-of-the-art (SOTA) method, Pix2Pix, in terms of PSNR and SSIM, while maintaining important image features while showing improved noise suppression.

After denoising, the next step is SAR image recognition. Vision Transformer (ViT) has shown high perfor-

DOI: 10.13164/re.2025.0429

mance on many recognition tasks such as Residual Neural Network (ResNet) or Visual Geometry Group (VGG) [9], but due to its nature relying on large datasets, it is not effective for SAR image analysis with a few labeled data. On the other hand, CNN-based recognition models are demonstrated to be able to identify SAR targets more robustly and efficiently even in a few data environments.

Therefore, this paper proposes a deep learning-based SAR recognition algorithm that integrates a CNN-based recognition model with a noise removal technique combining DnCNN and wavelet transform. Our approach is optimized for SAR image classification and low data SAR applications. Experimental results demonstrate that the proposed method not only achieves superior noise cancellation performance over SOTA method, but also improves SAR image recognition accuracy, demonstrating its potential in real-world SAR analysis applications.

Specifically, the main contributions of this paper are as follows. First, we propose a novel hybrid denoising approach that combines the strengths of DnCNN and wavelet transform, effectively integrating deep learning and signal processing techniques. Second, we demonstrate that our method outperforms conventional SOTA techniques such as Pix2Pix in terms of PSNR, SSIM, macro-average F1-score, and target recognition accuracy. Third, we show that the proposed framework maintains high performance even in low-SNR environments, making it especially practical for SAR applications.

The rest of this paper is organized as follows: Section 2 describes the proposed noise removal method, combining DnCNN and wavelet transform, and explains its advantages over conventional techniques. Section 3 presents the experimental setup, dataset details, and performance evaluation, comparing our approach with SOTA methods such as Pix2Pix. Section 4 discusses the results, including PSNR and SSIM metrics, and highlights the impact of our method on SAR image recognition. Finally, Section 5 provides the conclusions.

# 2. A Method to Enhance Image Recognition Based on Noise Removal for Navy SAR Radar

The SAR radar is a sensor that provides 2D image data using electromagnetic waves. The electromagnetic waves from the radar transmitter generate a radar image through the reception signal reflected from the target of interest. This signal is greatly affected by noise for various reasons such as maximum distance. Through this issue, the SAR radar recognition rate decreases. To improve this problem, the received radar image reduces the effect of noise through the DnCNN and wavelet transform.

## 2.1 DnCNN

The DnCNN technique used in this paper transforms the CNN network to fit image denoising and removes noise based on residual learning. Residual learning is a method of learning the residual, the difference between the input and output of a deep learning network. The DnCNN is trained to predict only noise from the output to the noisy input image. With the residual results obtained through this method, the DnCNN removes noise from the input image. First, a ground truth image $X$ is obtained, and then an Additive White Gaussian Noise (AWGN) noise is added to generate a noise image $Y$. A noise image $Y$ is prepared as the input of the CNN network, and noise $N$, which is the image pair, is prepared as the output. In the training stage, the noise removal ability is optimized by adjusting the weight of the network using the training data composed of the image pair of noise image $Y$ and noise $N$. It mainly uses a backpropagation algorithm, and an optimization algorithm is used to minimize the loss function. The learned DnCNN is tested by applying a new input image. Noise prediction is performed on the input image and an image with noise removed is generated by subtracting it from the original image.

The DnCNN method for residual learning uses the CNN method in Fig. 1. The SAR image noise is modeled as speckle noise in (1), which is multiplicative in nature and can be expressed as:

$$q = p \cdot n_{\mathrm{s}}, \quad n_{\mathrm{s}} \sim N(1, {\sigma_{\mathrm{s}}}^2), \tag{1}$$

$$q' = \log(q), \ q' = \log(p) + \log(n_{\mathrm{s}}). \tag{2}$$

Here, $p$ represents the clean SAR image, $q$ means the noisy SAR image and $n_{\mathrm{s}}$ is a multiplicative noise term with a mean of 1 and variance $\sigma_{\mathrm{s}}^2$. Speckle noise degrade image quality due to the coherent nature of SAR imaging systems. To solve this problem, the DnCNN employs residual learning to isolate and predict the noise component:

$$\omega' = R(q'; \theta), \tag{3}$$

$$\hat{p}' = q' - \hat{\omega}', \hat{p} = \exp(\hat{p}'). \tag{4}$$

Here, $R(q';\theta)$ represents the output of the DnCNN network, $\theta$ means the trainable parameters of the DnCNN network, $\hat{p}'$ is the restored log-domain SAR image and $\hat{p}$ is the final restored SAR image.

This process ensures that the DnCNN can effectively remove speckle noise for SAR target recognition. The DnCNN structure consists of $D$ layers, where $D = 29$. Each layer performs the following operations. The first convolution layer extracts feature from the input SAR image such

$$F_1 = \mathrm{ReLU}(\mathrm{Conv}_1(q'; W_1, b_1)) \tag{5}$$

where $\mathrm{Conv}_1$ is the first convolution layer, $W_1$ is weight and $b_1$ is bias. In intermediate layers, the batch normalization and ReLU activation is used for

$$F_k = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Conv}_k(F_{k-1}; W_k, b_k))) \tag{6}$$

where $k = 2,\ldots, D-1$ and BN means batch normalization layer. The final layer predicts the residual noise component

$\omega''$ is represented by

$$\hat{\omega}' = \mathrm{Conv}_D(F_{D-1}; W_D, b_D) \,. \tag{7}$$

These operations enable the DnCNN to focus on removing the noise for the SAR image.

Specifically, in the first layer, 64 feature maps are generated using 64 filters with a size of $3 \times 3 \times c$, and then the ReLU activation function is used as nonlinearity. Here, $c$ represents the number of image channels, $c$ = 1 for black

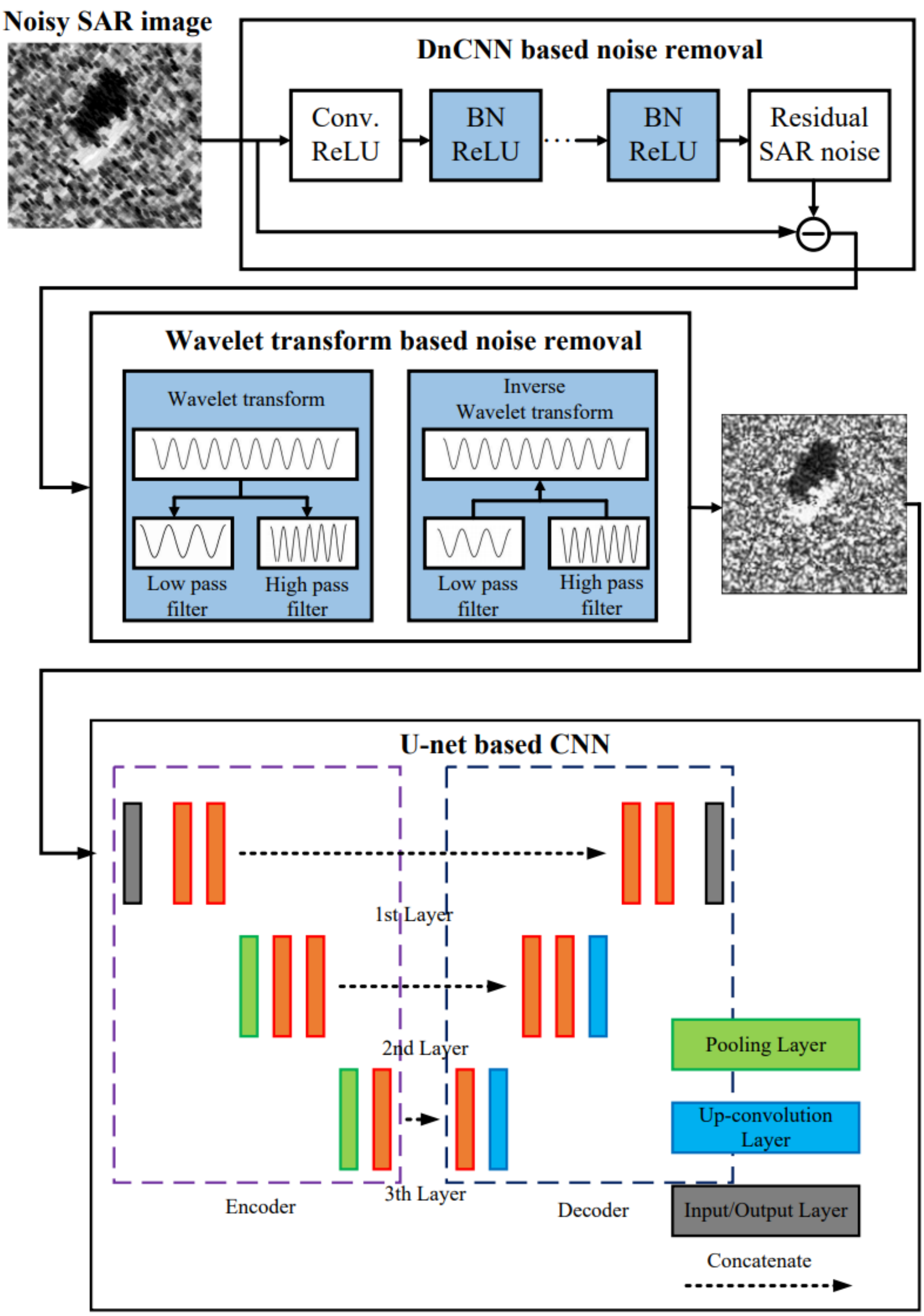



**Fig. 1.** The proposed structure.

| Layer | Details | Size |
|---|---|---|
| Input | SAR images | 1 x 128 x 128 |
| Conv_1 | 32 feature maps, kernel size = 3, ReLU | 32 x 128 x 128 |
| Pool_1 | Max pooling, kernel size = 2 | 32 x 64 x 64 |
| Conv_2 | 64 feature maps, kernel size = 3, ReLU | 64 x 64 x 64 |
| Pool_2 | Max pooling, kernel size=2 | 64 x 32 x 32 |
| Fc_3 | ReLU | 1 x 512 |
| Fc_4 | Dimension of CNN latent vector = 128 | 1 x 128 |
| Output | Num. of class = 8 | 1 x 8 |

(a) U-net encoder

| Layer | Details | Size |
|---|---|---|
| Fc_4_dec | Dimension of CNN latent vector = 128 | 1 x 512 |
| Fc_3_dec | ReLU | 64 x 32 x 32 |
| Pool_2_dec | Max pooling, kernel size = 2 | 64 x 64 x 64 |
| Conv_2_dec | 64 feature maps, kernel size = 3, ReLU | 32 x 64 x 64 |
| Pool_1_dec | Max pooling, kernel size = 2 | 32 x 128 x 128 |
| Conv_1_dec | 32 feature maps, kernel size = 3, ReLU | 1 x 128 x 128 |

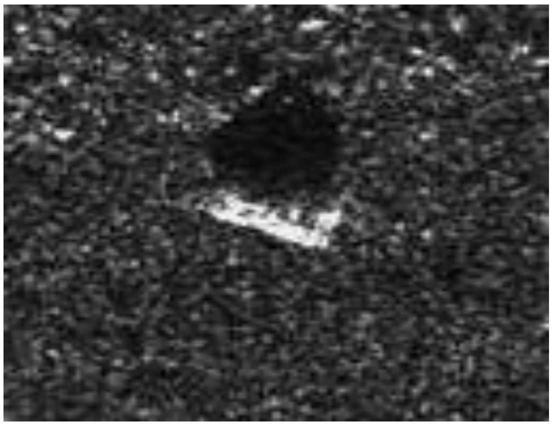
(b) U-net decoder

**Tab. 1.** Architecture of U-net based CNN.

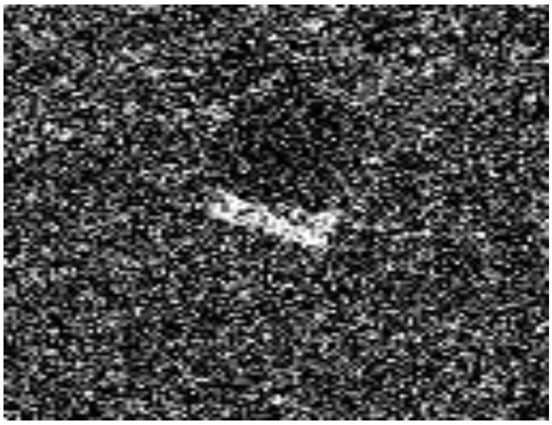
and white images and $c = 3$ for color images. In the second to $(D - 1)$, 64 filters with a size of $3 \times 3 \times 64$ are used, and batch normalization is added between the convolution and the ReLU. In the last layer, $c$ filters with a size of $3 \times 3 \times 64$ are used to reconstruct the output. In this paper, the number of network layers $D$ was set to 29. This utilized the results that the DnCNN noise removal network showed optimal performance in 20 to 30 layers in [6]. Therefore, it was determined that 29 layers provide sufficient learning capacity and provide a balance point to prevent overfitting.

## 2.2 Wavelet Transform

Wavelet transform is the method to decompose an arbitrary signal into functions defined as wavelets. Unlike Fourier transform decomposes a signal using cosine and sine functions that vibrate infinitely as basis functions, wavelet transform uses a function that has a limited vibration time as the basis function. It is widely used to remove noise by utilizing this wavelet transform [10]. To remove the noise of the radar signal in Fig. 1, first, we need to set the level $M$ and threshold values for wavelet transforming the signal after receiving the radar signal. The signal generated by wavelet transform is processed through the low pass filter and the high pass filter. We apply a thresholding technique that processes the result as zero for values lower than or equal to threshold. In this stage, the soft threshold was chosen because of its ability to gently suppress noise while maintaining critical high-frequency details. A global threshold approach was used at all levels to balance computational efficiency and performance. After that, the



(a) Original SAR data



(b) Noise-based SAR data

(c) The proposed SAR data

**Fig. 2.** Effect of noise removal through the proposed structure.

noise-removed wavelet transform result restores the signal through the inverse wavelet transform. As a result, the noise-removed radar signal is obtained. Compared to the conventional wavelet transform-only method mentioned in [4], the proposed approach combines the strengths of deep learning-based residual learning (DnCNN) and wavelet-based denoising. While wavelet transform effectively removes noise in the frequency domain, it struggles with complex spatial patterns.

The serial application of the DnCNN and the wavelet transform is designed to combine their respective strengths. The DnCNN removes the most of speckle noise while preserving critical high-frequency features through residual learning. Subsequently, the wavelet transform removes residual noise by removing the residual signal from the high-frequency band. For this reason, when the order of the steps is reversed or the steps are omitted, i.e., only wavelet transform, the performance is degraded in Figs. 3 and 4. In addition, the single DnCNN is difficult to remove residual noise on a fine scale, and only wavelet transform cannot effectively handle spatially dispersed noise.

### 2.3 U-Net Based CNN

The target is detected using a deep learning CNN model on the SAR image data from which noise has been removed. This CNN is based on U-net architecture [11], [12]. U-net is characterized by a U-shaped architecture that consists of a contracting path for feature extraction (encoder) followed by an expansive path for precise localization (decoder) as depicted in Fig. 1.

The encoder processes the input image and extracts features by using repeated stacks of convolutional layers with max pooling operations. Pooling operations reduce the spatial resolution of the image while capturing higher-level features. Each stack typically increases the number of filters, allowing for learning more complex features. We have a total of three levels of radar. The decoder aims to recover the spatial resolution while preserving the extracted features by using upsampling operations (like transposed convolution) to increase the resolution. Each upsampling step combines the upsampled feature map with a corresponding feature map from the contracting path (via skip connections). Skip connections directly provide detailed information from the earlier stages, helping the decoder accurately localize features. The final output is a reconstructed SAR image with noise removed, and the architecture allows for end-to-end learning for both classification and image reconstruction. This CNN-based approach effectively captures both spatial and hierarchical features, which are critical for accurate SAR target detection and noise removal. Table 1 shows the architecture of U-net based CNN. Figure 2 shows the effect of noise removal through the proposed structure.

## 3. Experimental Results

### 3.1 Quantitative Performance Analysis

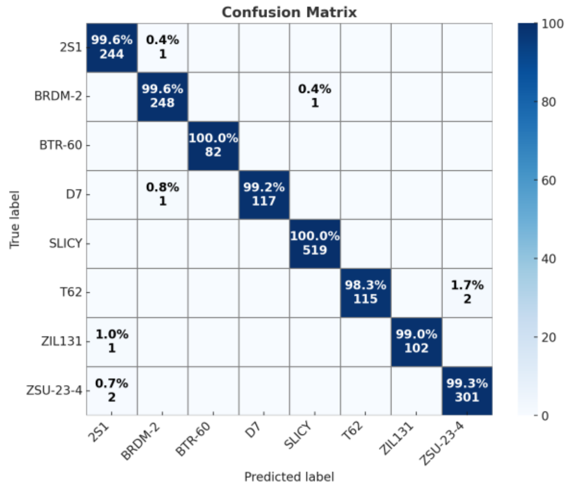
In this section, the results of the proposed model for the multi-class classification of SAR image radars are presented. The SAR data used the USA MSTAR data and the data obtained using 8 688 SAR images of 8 targets in the test set are utilized. For the simulation, the dataset was divided into 5 173 training sets, 1 778, and 1 737 test sets. When $N = 16$, the confusion matrix for the proposed model is shown in Fig. 3. The wavelet transform's level is set to 6. The test set is composed of 8 types including 2S1, BRDM-2, BTR-60 and D7, SLICY, T62, ZIL131, and ZSU-23-4. The SNR is set to 10 dB and reflects a noise condition that occur in the real environment and is a criterion frequently used in SAR image processing. This setting is used as a standardized method for evaluating the performance of models at a noise level and is considered an appropriate evaluation index considering the characteristics of SAR images in particular. The experiments were conducted using a high-performance computing system for efficient SAR image processing. Specifically, the system was equipped with an Intel Core i9-13900K CPU, 64 GB of DDR5 RAM, and an NVIDIA RTX 3090 GPU with 24 GB of VRAM, running on Windows 11. The software environment included Python 3.8 and PyTorch 1.12, with CUDA 11.3 utilized for GPU acceleration.

In Fig. 3(a), the average recognition rate of the original data is over 99.5%. Original SAR data refers to clean SAR images without added noise, obtained from the MSTAR dataset. This value represents the average of the recognition rates for each target type. When the SNR of the speckle noise in Fig. 3(b) is 10 dB, the average recognition rate of the confusion matrix decreased to approximately 94%, showing the impact of noise on the recognition performance. In Fig. 3(b), the noise-based SAR data refers to SAR images with speckle noise added at an SNR level of 10 dB. In Fig. 3(c), the confusion matrix of the wavelet transform-based method shows a performance degradation to approximately 81.6%. This indicates that wavelet transform alone is not effective in improving noise in SAR data. In Fig. 3(d), the reversed structure's average recognition rate is about 88.6%. The reversed structure refers to applying the wavelet transform first, followed by the DnCNN.

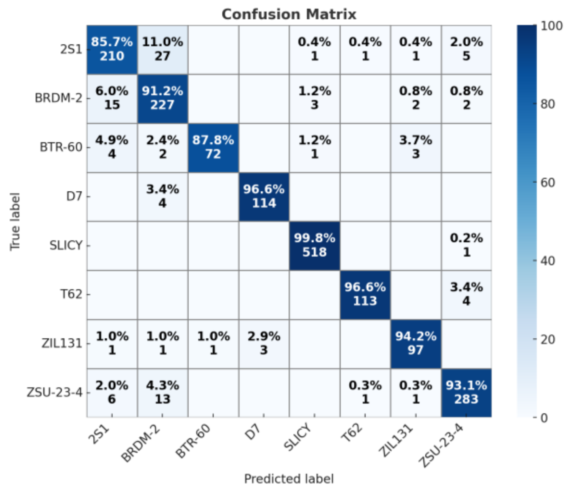
However, as shown in Fig. 3(e), the proposed structure improves the recognition rate to about 96%, demonstrating its effectiveness in reducing the influence of speckle noise and achieving a recognition rate close to that of the original data. In Fig. 3(e), the proposed SAR data refers to the denoised images produced by the proposed DnCNN and wavelet transform framework.

To provide a more comprehensive evaluation beyond recognition accuracy, we additionally computed the macro-average F1-score for each experimental condition. The F1-score reflects the harmonic mean of precision and recall and serves as a balanced metric to assess classification performance, especially in multi-class settings. The computed macro-average F1-scores for each method of U-net based CNN are as follows: 0.991 for the original SAR data, 0.868 for the noise-based SAR data, 0.789 for the wavelet transform-based SAR data, 0.841 for the reversed structure's SAR data, and 0.953 for the proposed SAR data. These results demonstrate that the proposed denoising and recognition framework significantly improves classification performance under noisy conditions, outperforming conventional and baseline methods in terms of both recognition accuracy and F1-score.

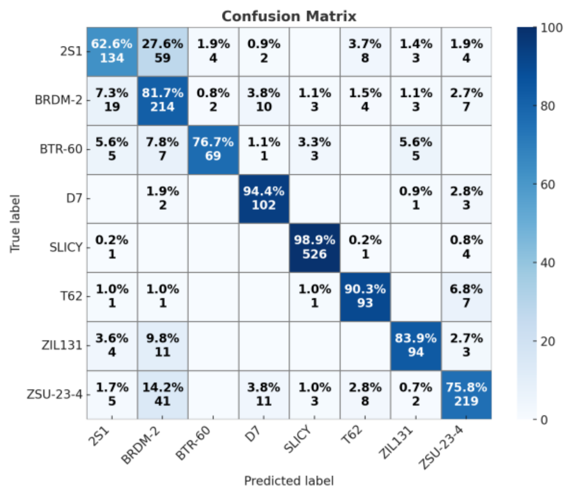
For comparison of the proposed model, the confusion matrix for the ViT model is shown in Fig. 4. The ViT is a transformer-based model proposed by [13] that applies self-attention mechanisms to image patches. For comparison purposes, we used a conventional ViT model with standard configurations, including patch size, transformer layers, and attention heads. The results of ViT were obtained using the same SAR dataset under identical noise conditions to ensure a fair comparison. Figure 4(a) shows



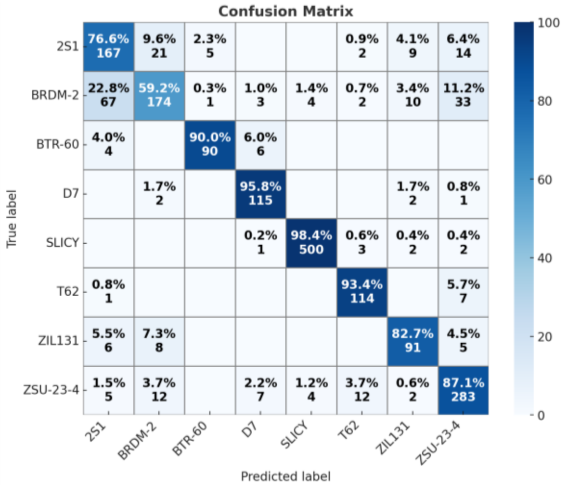

(a) Original SAR data



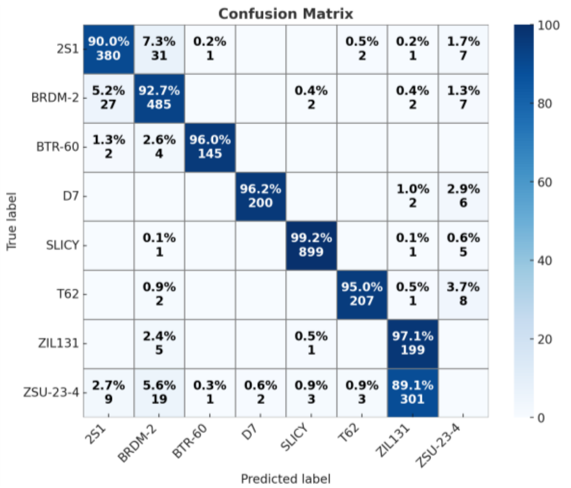

(b) Noise-based SAR data




(c) Wavelet transform-based SAR data




(d) Reversed structure's SAR data



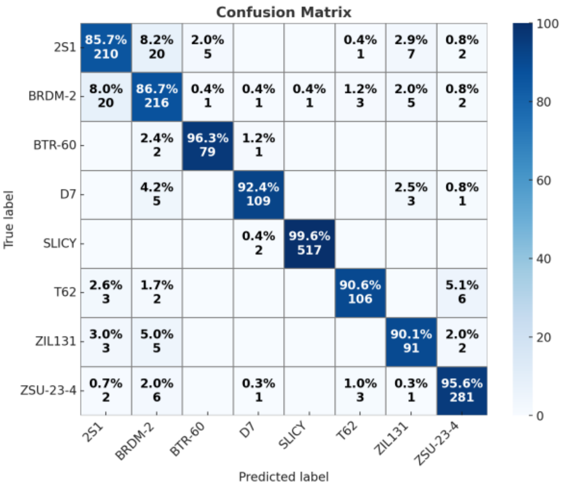

(e) The proposed SAR data

**Fig. 3.** U-net based CNN's confusion matrix.

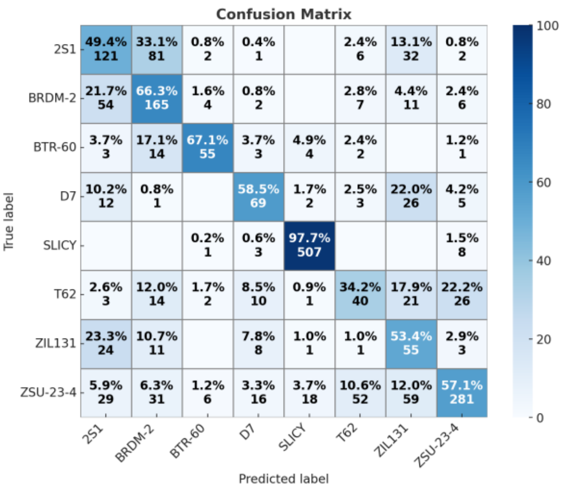
that the average recognition rate of the original data not affected by noise in the ViT model is about 91% or more. When the SNR of the noise in Fig. 4(b) is 10 dB, the average recognition rate of the confusion matrix decreases to about 58%. In Fig. 4(c), the confusion matrix of the wavelet transform-based method shows some improvement in performance to approximately 78.2%. This indicates that wavelet transform alone improves noise in SAR data but is still not sufficiently effective. In Fig. 4(d), the reversed structure's average recognition rate is about 87.4%. However, the average recognition rate of the network confusion matrix in Fig. 4(e) to which the proposed structure is applied has improved to about 92.4%.

The computed macro-average F1-scores for each method of the ViT model are as follows: 0.910 for the original SAR data, 0.567 for the noise-based SAR data, 0.765 for the wavelet transform-based SAR data, 0.859 for the reversed structure's SAR data, and 0.921 for the proposed SAR data. Moreover, the CNN-based model achieves better performance than the ViT-based model, particularly in cases with limited training data.

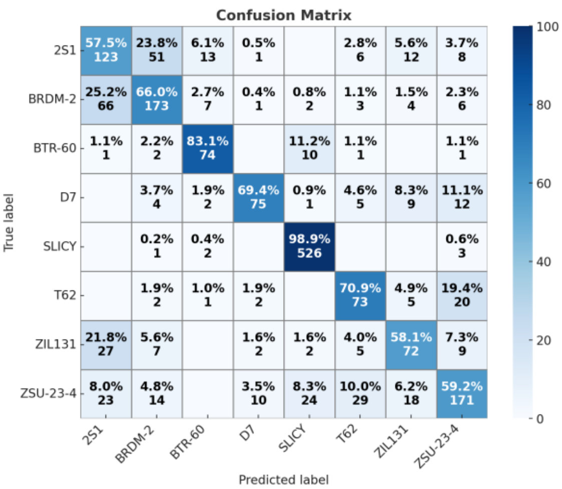
In addition to analyzing the confusion matrix, we evaluate the performance of the proposed noise cancellation method using standard metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index measurement (SSIM). As shown in Tab. 2, we confirm that the proposed method is superior to wavelet transformation-based noise cancellation via PSNR and SSIM. These results show that the combination of the DnCNN and wavelet transform-based noise cancellation effectively suppresses



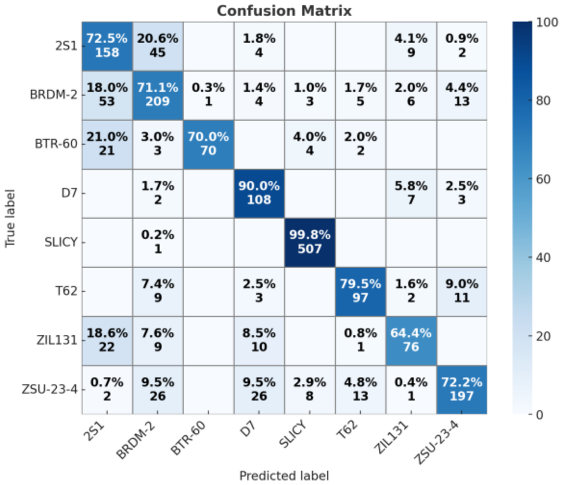

(a) Original SAR data



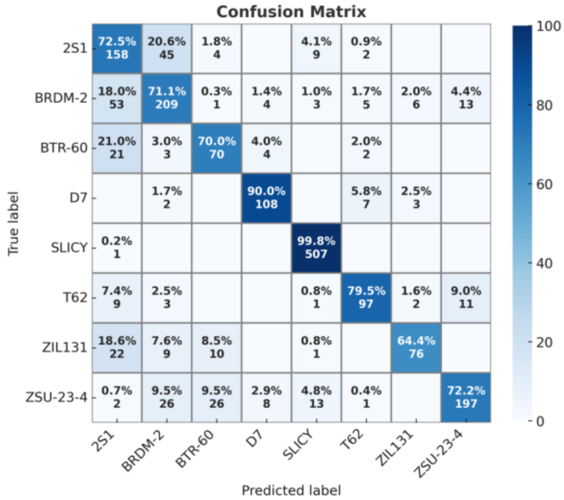

(b) Noise-based SAR data




(c) Wavelet transform-based SAR data




(d) Reversed structure's SAR data




(e) Conventional DnCNN, wavelet transform and ViT's SAR data

**Fig. 4.** Conventional ViT confusion matrix.

| Method | PSNR (dB) | SSIM |
|---|---|---|
| No method | 11.64 | 0.86 |
| Wavelet transform [4] | 14.10 | 0.83 |
| Pix2Pix [7] | 14.31 | 0.74 |
| Reversed structure | 15.54 | 0.84 |
| Proposed method | 16.88 | 0.89 |

**Tab. 2.** Performance evaluation of denoising methods using PSNR and SSIM metrics.

noise while preserving the structural details of the image. Significant improvements in the PSNR and SSIM indicate that the proposed method maintains more effectively similarity to the original image compared to other methods including SOTA.

## 3.2 Complexity Analysis

The computational complexity of the proposed method is influenced by the 29-layer DnCNN architecture and the wavelet transform. While the DnCNN provides effective noise removal through residual learning, its layer-by-layer operations involve convolution, ReLU activation, and batch normalization, contributing to a complexity of

$$O(L \cdot K^2 \cdot C_{\mathrm{in}} \cdot C_{\mathrm{out}} \cdot N^2) \tag{8}$$

where $L$ denotes the number of layers, $K$ represents the filter size, and $N$ is the size of a side of an image. $C_{\mathrm{in}}$ and $C_{\mathrm{out}}$ indicate the input and output channels, respectively. The wavelet transform adds an additional computational load with a complexity of

$$O(N^2). \tag{9}$$

For real-time SAR processing, the proposed method consists of 29-layer DnCNN and wavelet transforms, using (8) and (9) on a hardware platform with 1 TFLOPS processing to achieve a computational time of approximately 1.825 ms per $128 \times 128$ pixels of SAR image frame. This corresponds to a processing speed of over 500 frames per second, far exceeding the requirement of 30 frames per second for real-time tasks. Therefore, the proposed method is suitable for real-time SAR image processing applications.

### 3.3 Impact of Training Epochs on Model Performance

To further analyze the training and validation performance, we evaluated the effect of different training and test epochs on model convergence and performance. As shown in Fig. 5, increasing the epoch number initially improves generalization. And beyond a certain point in time, the model is optimized for the training data, stabilizing the performance of the test set. This observation indicates the importance of selecting the optimal epochs for training and conducting appropriate training and validation. Further details regarding the implementation of the proposed algorithm, including the source code and mathematical modeling, can be provided by the authors upon reasonable request.

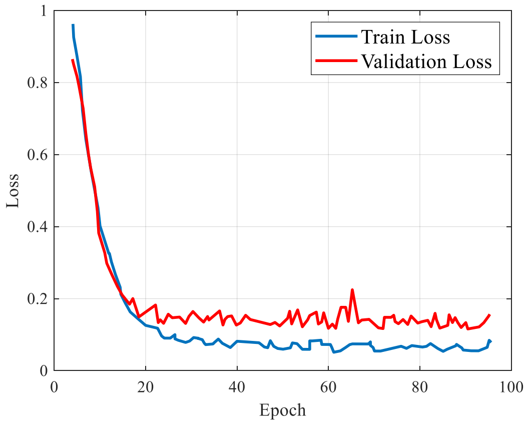



**Fig. 5.** Training and validation loss curve.

## 4. Conclusions

In this paper, the DnCNN and wavelet transform-based CNN recognition enhancement method for SAR radar was proposed. Through the proposed structure, in which the DnCNN algorithm, wavelet transform, and CNN are combined, the phenomenon of noise effects due to maximum distance or hardware limitations was effectively addressed. When the speckle noise level was set at 10 dB, the recognition performance was significantly improved, approaching the original SAR data's performance. Unlike ViT, which performs well with large-scale datasets, CNN is highly efficient even with smaller datasets, making them more suitable for SAR data, which often lacks the volume of labeled data required for ViT model. When using CNN, the average recognition rate of the original SAR data decreased from 99.5% to 94% in the presence of noise. However, applying the proposed CNN-based structure improves the recognition rate to 96%. Additionally, the method achieved a PSNR improvement from 2.2 dB to 7.4 dB and an SSIM enhancement from 0.3751 to 0.4416, indicating significant improvements in image quality. When using CNN, the macro-average F1-score decreased from 0.991 to 0.868 in the presence of noise. However, applying the proposed CNN-based structure improved the F1-score to 0.953. This demonstrates that CNNs, with their ability to handle noise effectively and their efficiency with limited data, are better suited for SAR radar applications where data volume is often constrained. Through this, the proposed model is expected to be a practical solution to effectively solve the noise problem frequently encountered in SAR radars, while also being more efficient for smaller datasets compared to ViT.

## Acknowledgments

This research was supported by the DGIST R&D Program of the Ministry of Science, ICT and Future Planning of Korea (25-IT-01 & 25-DPIC-4) and the 2025 academic research project of the Naval Institute for Ocean Research of the Republic of Korea Naval Academy. Youngdoo Choi and Geunhwan Kim are co-first authors. Bong-seok Kim and Sangdong Kim are co-corresponding authors.

## References

[1] ALBABA, A., BAUDUIN, M., SAKHNINI, A., et al. Sidelobes and ghost targets mitigation technique for high-resolution forward-looking MIMO-SAR. *IEEE Transactions on Radar Systems*, 2024, vol. 2, p. 237–250. DOI: 10.1109/TRS.2024.3366779

[2] ZHAI, Y., LIAO, J., SUN, B., et al. Dual consistency alignment based self-supervised learning for SAR target recognition with speckle noise resistance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, vol. 16, p. 3915–3928. DOI: 10.1109/JSTARS.2023.3267824

[3] AI, J., WANG, G., FAN, G., et al. A trilateral filter for video SAR speckle noise reduction. *IEEE Geoscience and Remote Sensing Letters*, 2023, vol. 19, p. 1–5. DOI: 10.1109/LGRS.2022.3174834

[4] CHOI, H., CHANG, J. Speckle noise reduction technique for SAR images using statistical characteristics of speckle noise and discrete wavelet transform. *Remote Sensing*, 2019, vol. 11, no. 10, p. 1–27. DOI: 10.3390/rs11101184

[5] MOHAN, E., RAJESH, A., SUNITHA, G., et al. A deep neural network learning-based speckle noise removal technique for enhancing the quality of synthetic-aperture radar images. *Concurrency and Computation: Practice and Experience*, 2021, vol. 33, no. 13, p. 1–12. DOI: 10.1002/cpe.6239

[6] ZHANG, K., ZUO, W., CHEN, Y., et al. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 2017, vol. 26, no. 7, p. 3142–3155. DOI: 10.1109/TIP.2017.2662206

[7] ZHAO, Y., CELIK, T., LIU, N., et al. A comparative analysis of GAN-based methods for SAR-to-optical image translation. *IEEE Geoscience and Remote Sensing Letters*, 2022, vol. 19, p. 1–5. DOI: 10.1109/LGRS.2022.3177001

[8] SUN, Y., YAN, K., LI, W. CycleGAN-based SAR-optical image fusion for target recognition. *Remote Sensing*, 2023, vol. 15, no. 23, p. 1–24. DOI: 10.3390/rs15235569

[9] HAQUE, M. F., LIM, H.-Y., KANG, D.-S. Object detection based on VGG with ResNet network. In *Proceedings of the 2019 International Conference on Electronics, Information, and Communication (ICEIC)*. Auckland (New Zealand), 2019, p. 1–3. DOI: 10.23919/ELINFOCOM.2019.8706476

[10] LI, W., ZHANG, Z., ZHANG, R. Newton time-reassigned multi-synchrosqueezing wavelet transform. *IEEE Signal Processing Letters*, 2024, vol. 31, p. 2390–2394. DOI: 10.1109/LSP.2024.3455990

[11] PHAM, T.-H., LI, X., NGUYEN, K.-D. SeUNet-Trans: A simple yet effective UNet-transformer model for medical image segmentation. *IEEE Access*, 2024, vol. 12, p. 122139–122154. DOI: 10.1109/ACCESS.2024.3451304

[12] WANG, L., SUN, Z., LI, Z., et al. TDWCNet: Triple UNet with dual-window convolution for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, vol. 62, p. 1–15. DOI: 10.1109/TGRS.2024.3444191

[13] DOSOVITSKIY, A., BEYER, L., KOLESNIKOV, A., et al. An image is worth 16×16 words: Transformers for image recognition at scale. In *Proceedings of the International Conference on Learning Representations (ICLR).* Virtual (Online), 2021. DOI: 10.48550/arXiv.2010.11929

## About the Authors ...

**Youngdoo CHOI** (co-first author) received the B.S. degree in Naval Science from the Korea Naval Academy in 2003, the M.S. degree in Electronic Engineering from Kyungpook National University in 2013, and the Ph.D. degree in Electronic Engineering from Kyungpook National University in 2016. Since 2021, he has been a Professor with the Department of Electronic Control Engineering at the Korea Naval Academy. His research interests include radar and sonar signal processing and underwater detection systems.

**Geunhwan KIM** (co-first author) was born in Daegu. He received the B.S. and M.S. degrees in Electronics Engineering and the Ph.D. degree in Electrical and Electronics Engineering from Kyungpook National University, Daegu, Republic of Korea, in 2015, 2017, and 2022, respectively. From 2022 to 2024, he held a Postdoctoral Researcher with the Department of Ocean System Engineering, Sejong University, Seoul, Republic of Korea. Since 2024, he has been with Changwon National University as an Assistant Professor. His research interests include underwater acoustic signal processing, high-duty-cycle sonar systems, and deep-learning applications for sonar systems.

**Bongseok KIM** (co-corresponding author) was born in Daegu. He received his B.S. in Electronics Engineering in 2006 and his M.S. and a Ph.D. in Information and Communications Engineering from Yeungnam University, South Korea, in 2009 and 2014, respectively. From 2014 to 2016, he held a postdoctoral position with the Daegu Gyeongbuk Institute of Science and Technology (DGIST), South Korea. Since 2016, he has been with DGIST as a senior research engineer. His current interests include multi-functional radar systems and radar signal processing. He is an IEEE Senior Member.

**Sangdong KIM** (co-corresponding author) was born in Seoul. He received his B.S. from Hanyang University in 2004, his M.S. from Han-yang University in 2006, and his Ph.D. from Kyungpook National University in 2018, all in Electronics Engineering. From 2015 to 2016, he was a visiting scholar at the University of Florida. From 2022 to 2023, he was a visiting scholar at Pennsylvania State University. Since 2006, he has been a principal researcher and adjunct associate professor with the DGIST. His research interests include super-resolution algorithms, automotive radar, and user authentication using vital sign radar. He is an IEEE Senior Member.