Journal of Beijing Institute of Technology, 2025, Vol. 34, No. 6

Image Denoising Based on Frequency Decomposition Generative Adversarial Network

Yulin Li[1], Mengmeng Zhang[1], Yunhao Gao[1], Wei Li[1], Gang Wang[2]✉

(1. School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China;
2. Quartermaster Engineering Technology Research Department, Systems Engineering Institute, Academy of Military Science, People's Liberation Army, Beijing 100010, China)

Abstract: In wireless communication scenarios, especially in low-bandwidth or noisy transmission conditions, image data is often degraded by interference during acquisition or transmission. To address this, we proposed Wasserstein frequency generative adversarial networks (WF-GAN), a frequency-aware denoising model based on wavelet transformation. By decomposing images into four frequency sub-bands, the model separates low-frequency contour information from high-frequency texture details. Contour guidance is applied to preserve structural integrity, while adversarial training enhances texture fidelity in the high-frequency bands. A joint loss function, incorporating frequency-domain loss and perceptual loss, is designed to reduce detail degradation during denoising. Experiments on public image datasets, with Gaussian noise applied to simulate wireless communication interference, demonstrate that WF-GAN consistently outperforms both traditional and deep learning-based denoising methods in terms of visual quality and quantitative metrics. These results highlight its potential for robust image processing in wireless communication systems.

Keywords: image denoising; deep learning; wavelet transform; edge guidance

1 Introduction

In real-world wireless communication scenarios, image data transmitted or captured is often degraded due to environmental interference, hardware limitations, and human-induced factors. These distortions typically manifest as noise, which not only compromises the visual quality of the images but also significantly hampers downstream tasks such as recognition, classification, and interpretation. Such distortions usually appear as noise, which not only reduces visual quality but also impairs downstream tasks such as recognition and classification. Such degradation may result in incorrect analysis outcomes in communication systems that rely on image data for decision-making [1].

Deep learning methods, with their ability to automatically extract complex features from large datasets, have made significant strides in the field of denoising. These models excel at preserving edges and fine details, significantly reducing edge blurring during the denoising process [2]. For example, the edge-focused image denoising (EFID) network proposed by Holla et al. [3] mitigates the loss of edge and texture information by using parallel processing and an edge recovery module, thereby enhancing the model's ability to represent details in denoising. These advancements have not only propelled the field forward but also set the stage for future work in image restoration tasks using deep learning.

Despite these advancements, deep learning

Manuscript received Jul. 24, 2025; revised Sep. 20, 2025; accepted Sep. 26, 2025. The associate editor coordinating the review of this manuscript was Dr. Lijuan Jia. The work was supported in part by the Beijing Natural Science Foundation (No. 4254072).

✉ Corresponding author. Email: 13910109005@126.com

DOI: 10.15918/j.jbit1004-0579.2025.042

techniques require adjustments to image size to improve efficiency and reduce hardware demands. Downsampling causes data loss, while upsampling may introduce artifacts, degrading the denoised image —especially for small objects or abstract textures where details are easily lost [4]. Liu et al. [5] were among the first to use wavelet transforms for adjusting feature map sizes, capitalizing on wavelet transforms' ability to capture both frequency and spatial information, thus retaining more image details while reducing model noise. Shahaf et al. [6] extended this approach by designing wavelet convolutions to increase the receptive field of the network, further improving the model's performance.

In the frequency domain, image information is usually divided into low- and high-frequency components. The former captures global structures and brightness, while the latter represents fine textures such as fur, hair, or architectural details [7]. However, most existing denoising methods focus primarily on modeling low-frequency regions, often neglecting the high-frequency components that are critical for restoring detailed textures and finer image information [6]. Jiang et al. [8] demonstrated that neural networks tend to prioritize fitting low-frequency signals first, leading to the preservation of global image structure while failing to restore the finer details effectively.

To this end, we proposed the Wasserstein frequency generative adversarial networks (WF-GAN) model, which is based on Wasserstein GAN with gradient penalty (WGAN-GP) and wavelet transforms. Unlike the original GAN, the input to the generator network in this case is not random noise; instead, it requires extracting sufficient meaningful information from the noisy images to guide the training process. The noise interference in the original images often manifests as abrupt changes, thus concentrating in the high-frequency domain. By leveraging wavelet transforms, we can obtain the low-frequency information of the input images. Additionally, by using contour detection to capture the overall outline, we can minimize the damage caused by high-frequency information to the contour details, thereby enhancing the quality of the generated images.

The contributions of this paper are as follows.

1) By using wavelet convolution to decompose the image, we employed a combination of high-pass and low-pass filters to separate the image into low-frequency and high-frequency components, enabling more refined denoising and ensuring that the final result appears more natural.

2) The wavelet transformation retains the overall structure of U-Net, whose superior feature extraction capabilities provide strong support for denoising. In the skip connection section, we designed a contour guidance module to enhance boundary information, further improving the quality of the denoised images.

3) To verify the reliability and effectiveness of the model, we conducted denoising tasks on multiple public datasets. Experimental results demonstrate the model's superiority in denoising performance, showing stronger noise suppression capabilities and better retention of image details compared to traditional methods.

2 Related Work

2.1 Traditional Method

Although traditional denoising methods have achieved limited progress in recent years, they are still actively studied due to their simplicity and interpretability [9]. These methods can be broadly classified into spatial domain and frequency domain approaches. Spatial domain methods analyze pixel neighborhoods to remove noise. Filters such as bilateral and guided filtering can preserve details and edges, with gradient-domain guided filtering offering further improvements in edge preservation [10]. Yin et al. [11] enhanced this by adjusting the regularization term for more generalized models. However, most spatial domain methods rely heavily on local neighborhood information, making them less effective when handling complex noise patterns or fine structures [12].

On the other hand, frequency domain methods take advantage of noise being concentrated in high-frequency regions [13]. By transforming images into the frequency domain, noise can be separated from useful signal. Hard thresholding removes coefficients above a certain threshold, preserving textures but potentially creating artifacts in high-frequency areas, while soft thresholding reduces these artifacts but introduces bias, which might lead to some loss of details [14]. To address these issues, Li et al. [15] proposed a soft thresholding method based on the non-subsampled shearlet transform (NSST), which efficiently distinguishes between noise and useful details, improving the overall denoising quality. Despite these improvements, frequency domain methods often struggle with adaptively balancing detail preservation and noise suppression, especially in scenarios where the noise distribution overlaps with fine image textures.

2.2 Deep Learning Method

Deep learning has achieved remarkable progress in image denoising owing to its powerful feature learning and generalization abilities [16]. A representative early work is denoising convolutional neural network (DnCNN) [17], which introduced residual learning by predicting the noise component instead of directly generating a clean image, thereby improving both accuracy and efficiency, but its design is mainly limited to additive white Gaussian noise. To address spatially varying noise, Zhang et al. [18] proposed fast and flexible denoising convolutional neural network (FFDNet), which incorporates noise level maps into the input to adaptively handle different noise intensities and improve robustness through orthogonal initialization; however, the method still depends on accurate noise estimation. More recently, Transformer-based models have emerged. Zamir et al. [19] proposed restoration transformer (Restormer), which employs multi-deformable Conv (DConv)-head transposed attention to capture global dependencies with reduced computational cost and achieves strong results in high-resolution image restoration. Similarly, Wang et al. [20] proposed U-shaped transformer (Uformer), which adopts a hierarchical U-shaped design with local window attention to balance local detail preservation and global context modeling, but requires careful parameter tuning to maintain efficiency.

In addition to spatial-domain approaches, several recent studies have explored frequency-domain processing for image denoising. Jiang et al. [21] proposed adaptive fusion dual-domain network (AFDN), a Fourier-based denoising network that applies high-pass filtering in the frequency domain to suppress noise components and enhance feature extraction. Fan et al. [22] proposed wavelet-domain frequency-mixing transformer (FMT), which employs a wavelet-based feature decomposition module to separate low- and high-frequency components and fuses them with spatial features to improve edge and texture preservation.

3 Method

This study improves upon the WGAN-GP model. First, the design of the discriminator combines conventional downsampling methods with residual blocks, ensuring a deep extraction of useful features during training while effectively avoiding information loss. Secondly, the structure of the generator retains the classic U-Net framework but replaces the downsampling and upsampling operations with wavelet transforms. Residual blocks are used in the feature learning process, ensuring the learning of image detail features and enhancing accuracy in denoising tasks. Additionally, an edge-guidance module is incorporated into the skip connections of the U-Net. This module enhances the accuracy of object contours, improving the precision of edge information in the generated images, thereby preventing edge blurring or loss. The overall structure of the Generator is shown in Fig. 1.

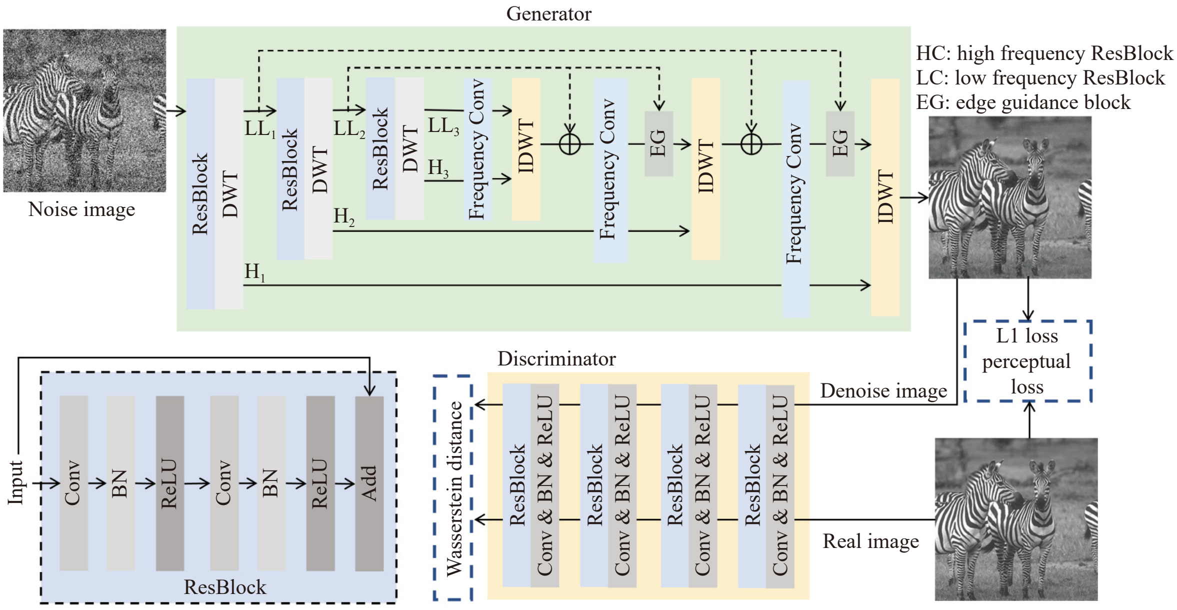

Fig. 1 The overall structure of the WF-GAN

3.1 Frequency Convolution Module

The effectiveness of wavelet transform largely depends on the choice of wavelet basis, as different bases decompose and reconstruct signals differently, directly affecting transformation performance [23]. The Daubechies wavelet is widely used due to its compact support, orthogonality, and excellent localization, which enable effective capture of local features in both spatial and frequency domains. The discrete wavelet transform (DWT) process can be expressed as follows

$$l_{j+1} = \sum_{k=0}^{K-1} l_j[2n-k]g_k \tag{1}$$

$$h_{j+1} = \sum_{k=0}^{K-1} l_j[2n-k]h_k \tag{2}$$

where l_j is the approximation at level j, n denotes the time/space index of the downsampled coefficients at scale j+1, and k is the filter tap index ranging from 0 to K–1, with K being the length of the analysis filters, and h_j is the detail component. The downsampling operator $[2n-k]$ ensures that the output length is half of the input, which is fundamental to the multiresolution property of wavelets. The filters g_k and h_k satisfy the quadrature mirror filter condition

$$\sum_{k=0}^{L-1} g_k g_{k-2n} = \delta[n] \tag{3}$$

$$h_k = (-1)^k g_{1-k} \tag{4}$$

where L is the filter length, which guarantees perfect reconstruction of the original signal. By iteratively applying Eq. (1) to the low-frequency component l_j, a multi-level decomposition is obtained. For two dimensional (2D) images, DWT is applied along rows and columns, resulting in four subbands at each level: low-low (LL), low-high (LH), high-low (HL), and high-high (HH). The LL subband captures the coarse structure, while the other three detail subbands preserve edges and textures.

The inverse DWT (IDWT) reconstructs the original image by upsampling and convolving the subbands with the synthesis filters g_k and h_k as follows

$$l_j[n] = \sum_{k=0}^{L-1} g_k\, l_{j+1}[n/2-k] + \sum_{k=0}^{L-1} h_k\, h_{j+1}[n/2-k] \tag{5}$$

This ensures that the global structure from LL and high-frequency details from LH, HL, HH are combined accurately. In our framework, feature maps from each subband are processed by convolutional layers and then fused via IDWT, as illustrated in Fig. 2, which preserves both coarse and fine image features for denoising.

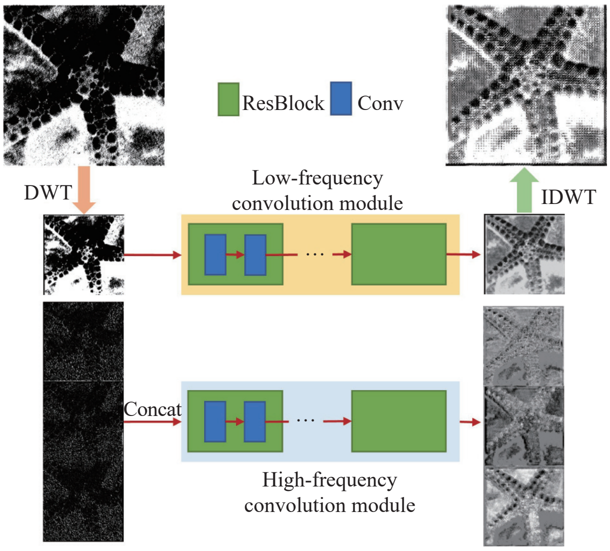

Fig. 2 Illustration of the frequency convolution module

3.2 Edge Guidance Module

During the image denoising process, although smooth regions are generally well restored, object contours —especially in areas with sharp color changes—often become blurred due to the overlap of high-frequency noise and edge information. To address this, the model extracts contour information from low-frequency feature maps as an auxiliary guiding signal. Specifically, given a low-frequency feature map $\boldsymbol{F}_{\text{low}} \in \mathbb{R}^{H \times W}$, the Sobel operator is applied to compute horizontal and vertical gradients as follows

$$\boldsymbol{G}_x = \boldsymbol{F}_{\text{low}} * \boldsymbol{S}_x \tag{6}$$

$$\boldsymbol{G}_y = \boldsymbol{F}_{\text{low}} * \boldsymbol{S}_y \tag{7}$$

where $*$ denotes convolution, and $\boldsymbol{S}_x$, $\boldsymbol{S}_y$ are the standard Sobel kernels

$$\boldsymbol{S}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{8}$$

$$\boldsymbol{S}_y = \boldsymbol{S}_x^{\mathrm{T}} \tag{9}$$

The edge magnitude map $\boldsymbol{E}$ is then computed as

$$\boldsymbol{E} = \sqrt{\boldsymbol{G}_x^2 + \boldsymbol{G}_y^2} \tag{10}$$

This edge map $\boldsymbol{E}$ serves as a guiding feature, ensuring that the network focuses on recovering object contours during training. Even when high-frequency components are affected by noise, the low-frequency contour information helps maintain the clarity of edges.

The model's overall structure is based on U-Net, with the edge-guided module functioning as part of the skip connection. The primary role of the skip connection is to combine high-resolution features from shallow layers with low-resolution features from deeper layers, preserving details and minimizing information loss. An attention mechanism is introduced to guide the learning of edge information, and convolution operations are applied after concatenation to further alleviate potential shifts. The module is illustrated as shown in Fig. 3.

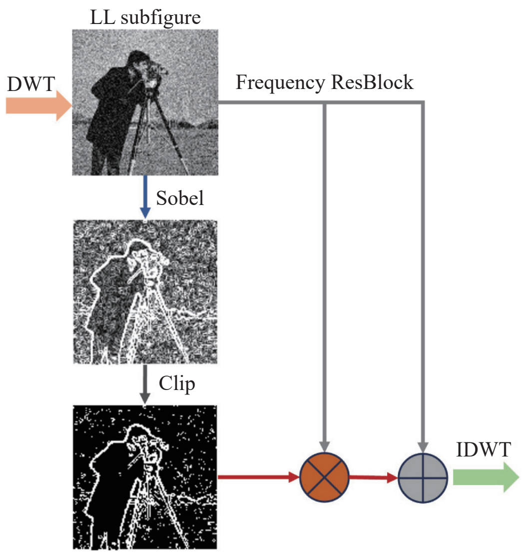

Fig. 3 Illustration of the edge guidance module

4 Experiments

4.1 Experimental Setup

In this study, we utilized a computer equipped with an NVIDIA real-time ray tracing (RTX) 4080 graphics processing unit (GPU), running on a Windows 11 operating system, and powered by an Intel Core i5 processor. The research was conducted using Python 3.10 and PyTorch 2.1.2. To evaluate the performance of the proposed model, we selected several widely recognized algorithms in the fields of machine learning and deep learning as benchmarks for comparison. WF-GAN was trained for 50000 iterations with a batch size of 12. The generator learning rate was 5×10^{-5}, the discriminator learning rate was 1×10^{-4}, and discriminator parameters were clipped to 0.01.

For model performance evaluation, we employed two main metrics: signal-to-noise ratio (SNR) and structural similarity index (SSIM) to assess denoising performance. The best results are highlighted in bold for ease of reference. Through these comparisons and evaluations, our objective is to comprehensively demonstrate the effectiveness and advantages of the proposed model.

4.2 Ablation Experiment

To validate the effectiveness of the designed modules, we performed ablation experiments on the Set12 data set with a noise level of 50. The experimental results show that incorporating wavelet transform significantly improves the model's denoising performance, demonstrating that a frequency-based approach can effectively extract more relevant features. Additionally, while the edge-guidance module does not show a notable improvement in peak signal-to-noise ratio (PSNR), it exhibits a significant enhancement in SSIM, indicating that this module effectively guides the restoration of the overall structure and contours of the image during the denoising process, particularly playing a key role in preserving detail and edge information. The experimental results are shown in Tab. 1.

4.3 Grayscale Image Denoising

In this study, the Berkeley segmentation dataset 400 (BSD400) was employed for training the grayscale image denoising model. The dataset was divided into training and validation sets, with approximately 70% of the images used for training and the remainder for validation. All images were cropped into patches of size 256×256 for model input. The performance of the trained model was evaluated on the Set12, BSD68, and Urban100 datasets as independent test sets. To simulate wireless communication interference at different intensity levels, Gaussian noise with standard deviations of 10, 25, and 50 was added to each dataset, enabling a comprehensive evaluation of the model's denoising performance in varying channel noise conditions. The experimental results, presented in Tab. 2, provide quantitative evaluation metrics for each dataset at different noise levels. The qualitative comparison results are shown in Fig. 4.

Among the compared methods, block matching 3D (BM3D) effectively preserves the overall image structure but is limited in restoring fine details, particularly at higher noise levels.

Tab. 1 Ablation experiment results

Feature extraction	Edge guidance	PSNR/SSIM (dB/%)
Conv	×	27.18/78.28
haar	×	27.15/79.19
db2	×	27.31/79.43
db6	×	27.64/80.13
db6	√	**27.92/81.11**

Tab. 2 Grayscale image denoising results in PSNR/SSIM

Dataset	σ	BM3D (dB/%)	DnCNN (dB/%)	NLRN (dB/%)	IRCNN (dB/%)	FFDNet (dB/%)	WF-GAN (dB/%)
	15	32.37/89.52	32.86/90.24	33.16/90.70	32.76/90.06	32.75/90.24	**33.21/91.23**
Set12	25	29.97/85.04	30.44/86.17	30.80/86.89	30.37/85.98	30.43/86.31	**31.00/87.52**
	50	26.72/76.76	27.18/78.28	27.64/79.80	27.12/78.04	27.32/78.99	**27.92/81.11**
	15	31.07/87.17	31.73/89.07	31.88/89.32	31.63/88.82	31.63/89.02	**31.95/89.86**
BSD68	25	28.57/80.13	29.23/82.79	29.41/83.31	29.15/82.48	29.19/82.88	**29.56/84.02**
	50	25.60/68.64	26.23/71.89	26.47/72.98	26.19/71.69	26.29/72.39	**26.67/73.96**
	15	32.35/92.20	32.64/92.46	33.45/93.54	32.46/92.36	32.40/92.65	**33.58/93.96**
Urban100	25	29.70/87.77	29.95/87.81	30.94/90.18	29.80/88.31	29.90/89.79	**31.23/91.41**
	50	25.95/77.91	26.26/78.56	27.49/82.79	26.22/79.18	26.50/80.47	**27.83/84.83**

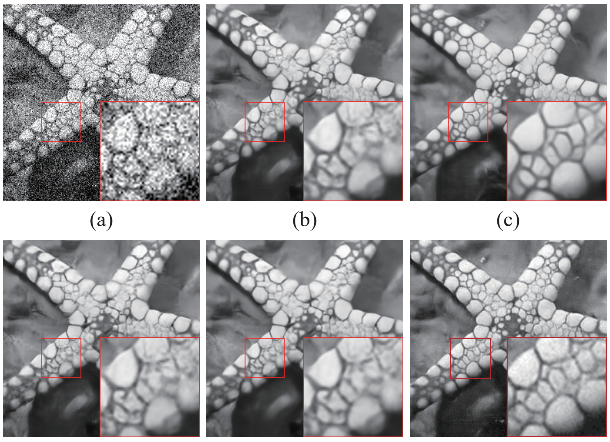

Fig. 4 Grayscale denoising results of various models: (a) noise image; (b) DnCNN; (c) NLRN; (d) IRCNN; (e) FFD-Net; (f) WF-GAN

DnCNN and FFDNet achieve a moderate balance, providing reasonable recovery of both edges and textures. Non-local recurrent network (NLRN) emphasizes edge and contour sharpness, sometimes producing details that exceed those in the original images, yet this often introduces artifacts and reduces global consistency. CNN denoiser prior for image restoration (IRCNN) generally maintains the global structure but is less capable of capturing high-frequency details. In contrast, the proposed WF-GAN demonstrates superior performance, consistently recovering both fine details and overall image structures, yielding denoised results that are visually more faithful and quantitatively stronger.

4.4 Color Image Denoising

For the color image denoising task, the Berkeley dataset was employed for training the model, with the training/validation split and patch size following the same procedure as in the grayscale experiments. The performance of the trained model was evaluated on four widely-used public datasets: color Berkeley software distribution 68 (CBSD68), Kodak24, McMaster, and Urban100. Gaussian noise with standard deviations of 15, 25, and 50 was added to each dataset, as in the grayscale image processing task. The results are summarized in Tab. 3, comparing our model's performance with other algorithms at different noise levels. For qualitative comparison, image 13 from the McMaster dataset is shown in Fig. 5.

In the task of color image denoising, the compared methods exhibit distinct behaviors. DnCNN and IRCNN deliver comparable yet relatively basic results, indicating limited capability in modeling multi-channel noise. FFDNet shows certain stability at moderate to high noise levels, but its performance in color fidelity and fine detail preservation remains insufficient. Deep CNN with batch renormalization (BRDNet) achieves strong results at low noise levels, particularly on Urban100, though its advantage diminishes rapidly as noise intensity increases. Residual non-local attention networks (RNAN) demonstrates potential in specific high-noise con-

Tab. 3 Color image denoising results in PSNR/SSIM

Dataset	σ	RNAN (dB/%)	BRDNet (dB/%)	DnCNN (dB/%)	IRCNN (dB/%)	FFDNet (dB/%)	WF-GAN (dB/%)
CBSD68	15	–	34.10/92.91	33.89/92.90	33.87/92.85	33.87/92.90	**34.26/93.10**
	25	–	31.43/88.47	31.23/88.30	31.18/88.24	31.21/88.21	**31.59/88.79**
	50	28.27/80.18	28.16/79.42	27.92/78.96	27.88/78.98	27.87/83.96	**28.48/80.96**
Kodak24	15	–	34.88/92.49	34.48/92.09	34.69/92.09	34.63/92.24	**35.08/92.92**
	25	–	32.41/**88.56**	32.03/87.75	32.15/87.79	32.13/87.91	**32.74/88.56**
	50	29.58/81.18	29.22/80.40	28.85/79.17	28.94/79.43	28.98/79.52	**29.80/81.82**
McMaster	15	–	35.08/92.69	33.44/90.35	34.58/91,95	34.66/92.16	**35.29/92.88**
	25	–	32.75/89.43	31.51/86.94	32.18/88.18	32.35/88.61	**33.07/90.13**
	50	29.72/83.29	29.52/82.65	28.61/79.86	28.93/80.69	29.18/81.49	**29.90/83.92**
Urban100	15	–	**34.42**/94.62	32.98/93.14	33.78/93.14	33.83/94.18	34.19/**94.93**
	25	–	31.99/91.94	30.81/90.15	31.20/90.88	31.40/91.20	**32.22/92.31**
	50	29.08/87.03	28.56/85.77	27.59/83.31	27.70/83.96	28.05/84.76	**29.29/87.84**

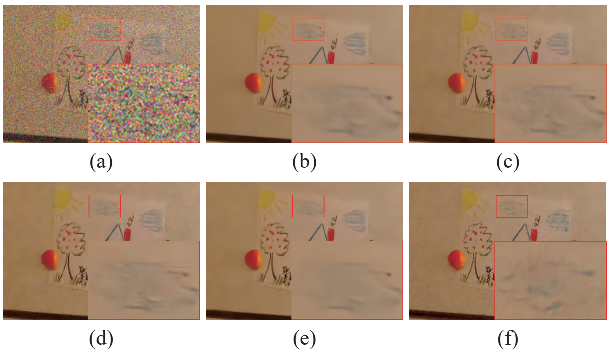

Fig. 5 Color denoising results of various models: (a) noise image; (b) RNAN; (c) BRDNet; (d) IRCNN; (e) FFD-Net; (f) WF-GAN

ditions but lacks overall consistency. Compared to existing methods, WF-GAN achieves similar PSNR but significantly higher SSIM, indicating better preservation of structural details and perceptual quality. Its advantage is particularly evident in both color and grayscale images with complex textures, though there remains room for improvement in extremely high-noise scenarios.

5 Conclusion

In this paper, we proposed the WF-GAN model, designed for effective denoising of noisy images in wireless communication environments from a frequency perspective. By decomposing and analyzing different frequency bands, we can perform feature learning tailored to the characteristics of each band, significantly improving denoising performance. The wavelet decomposition technique allows for the full utilization of information from different frequency bands while reducing noise interference with the overall structure, thereby enhancing the clarity of the generated images. Our model also incorporates an edge-guidance module, which extracts high-frequency edge information hidden beneath severe noise interference using low-frequency signals. This further improves the detail quality of the denoised images. Additionally, we employed frequency loss and edge loss to mitigate potential detail texture loss caused by discriminator loss and pixel-wise loss. Experimental results demonstrate that WF-GAN achieves excellent denoising performance, showcasing its broad applicability in wireless communication image processing.

While the proposed WF-GAN demonstrates strong denoising performance, several avenues for future improvement remain. Key research directions include reducing the impact of noise in high-frequency components during wavelet decomposition and developing more adaptive feature learning strategies that fully exploit the characteristics of different frequency bands. Our future work will focus on addressing these challenges to further enhance the model's denoising capability and robustness.

References:

[1] R. S. Jebur, M. H. B. M. Zabil, D. A. Hammood, and L. K. Cheng,"A comprehensive review of image denoising in deep learning," *Multimedia Tools and Applications*, vol. 83, no. 20, pp. 58181-58199, 2024.

[2] S. Izadi, D. Sutton, and G. Hamarneh,"Image denoising in the deep learning era," *Artificial Intelligence Review*, vol. 56, no. 7, pp. 5929-5974, 2023.

[3] S. Holla, N. Park, and B. Lee,"EFID: Edge-focused image denoising using a convolutional neural network," *IEEE Access*, vol. 11, pp. 9613-9626, 2023.

[4] N. Liu, L. Wu, J. Wang, H. Wu, J. Gao, and D. Wang,"Seismic data reconstruction via wavelet-based residual deep learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022.

[5] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-CNN for image restoration," 2018. https://arxiv.org/abs/1805.07071.

[6] S. E. Finder, R. Amoyal, E. Treister, and O. Freifeld, "Wavelet convolutions for large receptive fields," 2024. https://arxiv.org/abs/2407.05848.

[7] H. Yu, N. Zheng, M. Zhou, J. Huang, Z. Xiao, and F. Zhao, "Frequency and spatial dual guidance for image dehazing," in *European Conference on Computer Vision*, Cham, Switzerland, pp. 181-198, 2022.

[8] L. Jiang, B. Dai, W. Wu, and C. C. Loy, "Focal frequency loss for image reconstruction and synthesis," 2021. https://arxiv.org/abs/2012.12821.

[9] A. Sharma and V. Chaurasia,"MRI denoising using advanced NLM filtering with non-subsampled shearlet transform," *Signal, Image and Video Pro-*

cessing, vol. 15, no. 6, pp. 1331-1339, 2021.

[10] B. Goyal, A. Dogra, S. Agrawal, and B. Sohi, "Two-dimensional gray scale image denoising via morphological operations in NSST domain & bitonic filtering," *Future Generation Computer Systems*, vol. 82, pp. 158-175, 2018.

[11] H. Yin, Y. Gong, and G. Qiu, "Guided filter bank," in *Intelligent Computing: Proceedings of the 2021 Computing Conference*, pp. 783-792, 2021. https://doi.org/10.1007/978-3-030-80119-9_50.

[12] Z. Yan, W. Xu, C. Ai, and Y. Geng, "Parametric study on pump noise processing method of continuous wave mud pulse signal based on dual-sensor," *Journal of Petroleum Science and Engineering*, vol. 178, pp. 987-998, 2019.

[13] Z. Li, H. Liu, L. Cheng, and X. Jia, "Image denoising algorithm based on gradient domain guided filtering and NSST," *IEEE Access*, vol. 11, pp. 11923-11933, 2023.

[14] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142-3155, 2017.

[15] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608-4622, 2018.

[16] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 5718-5729, 2022.

[17] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general U-shaped transformer for image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 17683-17693, 2022.

[18] Z. Jiang, Z. Xue, J. Wang, Y. Hu, and Q. Zheng, "Single image denoising based on adaptive fusion dual-domain network," *IET Image Processing*, vol. 18, pp. 561-571, 2024.

[19] H. Fan, M. Lu, X. Zhang, Z. Ma, P. Liu, C. Yang, L. Kong, R. Chen, and X. Cao, "Wavelet-domain frequency-mixing transformer unfolding network for low-dose computed tomography image denoising," *Quantitative Imaging in Medicine and Surgery*, vol. 15, no. 8, pp. 7419, 2025.

[20] A. Halidou, Y. Mohamadou, A. A. A. Ari, and E. J. G. Zacko, "Review of wavelet denoising algorithms," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 41539-41569, 2023.

[21] L. Cheng and P. Chen, "Remote sensing image denoising based on Gaussian curvature and shearlet transform," *IEEE Access*, vol. 11, pp. 97716-97725, 2023.

[22] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided CNN for image denoising," *Neural Networks*, vol. 124, pp. 117-129, 2020.

[23] L. Li, L. Wang, Z. Jia, Y. Si, J. Yang, and N. K. Kasabov, "A practical medical image enhancement algorithm based on nonsubsampled contourlet transform," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 5, pp. 1046-1056, 2019.

Yulin Li received the B.S. degree in communication engineering from Tianjin University, Tianjin, China in 2021, and the M.S. degree in communication engineering from Beijing University of Chemical Technology, Beijing, China in 2024. He is currently pursuing the Ph.D. degree with School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His research interests include deep learning, image denoising, and hyperspectral image generation.

Mengmeng Zhang received her B.S. degree in computer science and technology from the Qingdao University of Science and Technology, Qingdao, China, in 2014, and the Ph.D. degree in control science and engineering from Beijing University of Chemical Technology, Beijing, China, in 2019. She is currently an Assistant Professor with School of Information and Electronics, Beijing Institute of Technology, Beijing, China. Her research interests include remote sensing image processing and pattern recognition.

Yunhao Gao received the B.S. and M.S. degrees from College of Information Science and Engineering, Ocean University of China, Qingdao, China, in 2017 and 2020, respectively, and the Ph.D. degree from School of Information and Electronics, Beijing Institute of Technology, Beijing, China, in 2024. He is currently a Postdoctoral Researcher with School of Information and Electronics, Beijing Institute of Technology. His current research interests include multisource remote sensing collaboration and interpretation.

Wei Li received the B.E. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2007, the M.S. degree in information science and technology from Sun Yat-Sen University, Guangzhou, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2012. Subsequently, he spent 1 year as a Postdoctoral Researcher at the University of California, Davis, CA, USA. He is currently a Professor with School of Information and Electronics, Beijing Institute of Technology. His research interests include hyperspectral image analysis, pattern recognition, and target detection.

Gang Wang is currently with the Quartermaster Engineering Technology Research Department, Systems Engineering Institute, Academy of Military Science, People's Liberation Army. His research interests include remote sensing image processing, deep learning, and related applications.