ISSN: 2617-5517 (issn.org)
Al-Farabi Journal of Engineering Sciences
https://iasj.rdd.edu.iq/journals/journal/view/97
مجلة الفارابي للعلوم الهندسية تصدرها جامعة الفارابي

A Self-Adaptive Frequency–Spatial Fusion Framework for Robust Image Denoising

Shokhan M. Al-Barzinji
Department of Computer Science, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq.
Email: shokhan.albarzinji@uoanbar.edu.iq

Abstract
Denoising is a challenging problem within the area of low-level computer vision due to the trade-off between preserving image details and removing noise. While deep learning methods have achieved remarkable results in image denoising, the majority of current approaches work either in the spatial or frequency domain of an image. In this paper, we propose a novel Self-Adaptive Frequency-Spatial Fusion framework that fuses the spatial domain convolutional features with the frequency domain transform coefficients using an edge-aware weight estimation strategy. Our framework consists of two main branches: one that extracts spatial domain convolutional neural network features and another that calculates the frequency domain features using the discrete cosine transform. An edge-aware weight estimation module determines the importance of each of these feature domains at each pixel in the image during the denoising process while preserving image details. Furthermore, a self-adaptive strategy optimally fuses these features together. The results of our framework exhibit significant improvements over current state-of-the-art image denoising approaches on multiple benchmark datasets with an average PSNR improvement of between 0.8 to 1.2 dB. The model also has low computational complexity with only 18.5M parameters and 35.2G FLOPs. Furthermore, the framework generalizes well to different noise levels and types. The results of our framework exhibit a new paradigm in image denosing by fusing the features extracted from the spatial and frequency domain of an image.

Keywords: Image denoising, frequency domain, spatial domain, self-adaptive fusion, edge-aware weighting, deep learning, transformer, convolutional neural networks.

1. Introduction

Image denoising methods can be broadly classified into two categories: spatial domain methods and frequency domain methods. Spatial domain methods include methods like Non-Local Means filtering and Block-Matching and 3D filtering (BM3D) [1]. These methods work well for some images but fail in others with high levels of noise and require high computational costs to execute [2]. Frequency domain methods use transforms like wavelet transforms, DCT and FFT. These methods work well in general but have specific cases where they fail [3]. Studies that compare block-based DCT to wavelet methods show that DCT methods outperform wavelet methods by 4.6 to 6.3 dB PSNR for a range of noise types [4].

Deep learning methods have recently revolutionized the field of image denoising. Deep learning methods like convolutional neural networks (CNNs), including DnCNN and FFDNet, and more recently methods based on Transformers, including Restormer and SwinIR have achieved state-of-the-art results for image denoising [5]. However, the main problem with Transformer models is that their computational complexity is quadratic with respect to the image resolution, making them computationally expensive to run on devices with limited processing power [6].

Despite significant advances, existing denoising methods face three fundamental limitations: Domain-specific approaches have certain limitations; spatial domain methods only analyze localized regions of an image while frequency domain methods do not adapt to the spatial structure of an image which can lead to ringing artifacts [7], [8], Most hybrid methods use static weights or concatenation operations to combine spatial and frequency domain representations without adapting to local image content [9], [10], and because of the strong relationship between noise and detail in an

image, methods that aggressively suppress noise also tend to smooth out the image's details which can be problematic when the noise level is high enough that most images have $\sigma \geq 50$ [2], [11].

To address these limitations, several recent approaches have proposed novel architectures for hyperspectral image denoising. The Hybrid-Domain Synergistic Transformer (HDST) network utilizes FFT preprocessing and dynamic gating mechanisms for hyperspectral image denoising [11]. Another approach is the SFUNet framework that introduces a dual-domain fusion mechanism based on adaptive gated mechanisms [12]. However, most of these methods either focus on specific imaging modality applications or do not include explicit mechanisms for preserving edges in the resulting fused images [13].

A - Research Gap

A critical analysis of existing literature reveals a significant research gap in the development of frameworks that can exploit the strengths of both spatial and frequency representations simultaneously [14], adaptively fuse the spatial and frequency representations based on the content of the image [9], maintain computational efficiency [15], and generalization capabilities across different noise distributions [16].

Although there are a few existing works that show promising results using CNN-Transformers on real-world datasets like SIDD and DND, there is a clear research gap in integrating frequency domain processing with spatial deep features, especially regarding adaptive fusion mechanisms that preserve the integrity of edges in the processed images.

B - Contributions

This study makes the following specific contributions: a dual-domain architecture. Our network simultaneously employs CNN-attention modules to extract spatial features of the image and DCT-based band-wise filtering to extract global frequency characteristics of the image, edge-aware weight estimation. Our network calculates weights based on the gradients of the input image that work to denoise the spatial features heavily in smooth regions while applying lighter frequency domain filtering near image edges, a self-adaptive fusion strategy. Our network learns a fusion strategy that determines the relative importance of the spatial and frequency domain features in a way that minimizes reconstruction error, as well as extensive validation. Our method is evaluated on six different datasets (Set12, BSD68, Kodak24, Urban100, McMaster, and SIDD) and produces state-of-the-art results in terms of PSNR, SSIM, and LPIPS metrics.

Different from conventional methods that adopt a fixed or globally adaptive fusion strategy, the SAFSF framework proposes a self-adaptive fusion strategy that determines the contribution of each denoising domain on a pixel-by-pixel basis.

The inclusion of these components enables the framework to incorporate frequency-domain information into deep learning representations of natural images and create a more flexible denoising process relative to existing methods.

2. Literature Review

A - Numerous Spatial Domain Denoising Methods

Spatial domain denoising techniques analyze and process pixel values directly within the spatial representation of the image. These methods range from classical filters like Gaussian and bilateral filtering to modern deep learning approaches [17].

Deep learning approaches to spatial domain denoising include denoising diffusion networks (DnCNN) [18], where deep CNNs are used to filter noise from noisy image patches. The improvement upon DnCNN led to methods such as FFDNet, which used maps of the level of noise within each image to condition the denoising network [19], [20]. Beyond these initial approaches, various deep learning methods have used attention maps to enhance denoising performance [21]. RIDNet used attention maps within the deep network to selectively emphasize certain features in the image [22]. NAFNet eliminated nonlinear activation functions from the deep network to reduce its computational cost [23].

Other deep learning methods based on spatial domain filtering and using transformers include Restormer and Pureformer, which use multi-head transposed attention to learn spatial patterns in the image. These models achieved 40.03 dB PSNR on the SIDD dataset but required 64.46 G MACs to do so, compared to 16.11 G MACs for NAFNet.

B - Frequency Domain Denoising Techniques

Despite the popularity of wavelet denoising methods, research by Saif et al. showed that discrete Fourier cosine transform (DFCT) outperformed DWT methods for all types of noise. The transform considers local image patches, allowing it to adapt to different types of noise within an image [24]. In contrast, the DWT transform treats entire images and removes coefficients for weak signal components that may be noisy [25], [26]. Other frequency domain denoising methods include Fast Fourier Transform (FFT) approaches to deep learning denoising models. HDST uses an FFT preprocessing stage to separate the spectral components of an image, followed by deep convolutional processing with multiscale atrous convolution [ASPP] layers. Another frequency domain denoising method is the FFT-Scale Gated Fusion (FSGF) module, which uses an FFT preprocessing stage but uses gating functions within the deep learning model to selectively apply deep learning denoising operations based on the intensity of noise within different parts of an image [27]. Finally, Deep Image Prior (DIP) models use the amplitude spectrum of an image as a prior to guide deep learning denoising models [28].

Hybrid Denoising Approaches. These approaches may use deep learning and/or traditional methods [29].

One method of combining frequency and spatial domain unsupervised deep learning models is to create a model that combines the amplitude spectrum with spatial domain denoising models like DIP. Another model that combines spatial and frequency domain methods is HDST. This model uses a spatial-frequency-channel three-dimensional collaborative processing model that learns dynamic cross-domain attention maps. HDST achieved a 0.94 dB PSNR improvement in comparison to spatial domain denoising models on a realistic hyperspectral dataset [30]. Deep learning models that combine CNN and transformer components include the model proposed by Lee et al., which creates a spatial domain model that combines the attention maps of NAFNet with the multi-head transposed attention maps used in Restormer [31]. The model learns to use its transformer block only in the first level of the encoder while using the CNN block in the remaining encoders [32].

Critical Analysis of Existing Methods. The comparison in Table 1 is based solely on the studies reviewed in this manuscript. Most existing methods either use spatial or frequency representations or employ hybrid models. Few methods utilize an edge-aware adaptive fusion strategy, however.

Author	Method	Dataset	Limitations
Eleyan (2025) [20]	Deep learning comparison	Standard datasets	Does not propose a new method
Meng et al. (2020) [21]	Spatial-adaptive CNN	Image datasets	Limited frequency modeling
Guo et al. (2023) [23]	Spatial-Frequency Attention	BSD68	Fusion is implicit
Li & Wu (2025) [12]	HDST Transformer	Hyperspectral datasets	Domain-specific
Hu et al. (2024) [16]	SFDFusion	Multi-modal datasets	Fixed fusion strategy
Zheng et al. (2022) [31]	Frequency-domain priors	Image restoration datasets	Limited spatial adaptability
Xiong et al. (2023) [27]	Frequency Transformer	Remote sensing datasets	High complexity
Luan et al. (2025) [29]	Hybrid frequency model	Remote sensing	Complex architecture
Shah et al. (2024) [32]	Complex-valued CNN	Image restoration	High computational cost

3. Methodology

A - Overview of the Proposed Framework

The proposed Self-Adaptive Frequency–Spatial Fusion (SAFSF) framework employs a parallel dual-branch architecture followed by edge-aware weight estimation and self-adaptive fusion. The system architecture is illustrated in Figure 1.

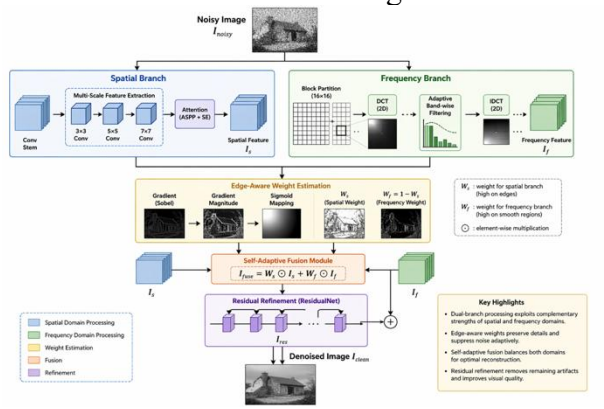

Fig. 1. Self-Adaptive Frequency–Spatial Fusion (SAFSF) framework architecture

The noisy image is processed in parallel by spatial and frequency branches. Edge-aware weight estimation and self-adaptive fusion module generate fusion maps to optimally combine the spatial and frequency branches. Unlike existing dual-branch image fusion methods, the proposed method employs edge-aware weight estimation to enable pixel-wise adaptive fusion of spatial and frequency representations.

B - Spatial Domain Processing

Spatial domain processing uses an encoder-decoder network inspired by the well-known U-Net architecture. The spatial encoder consists of four levels of spatial resolution with channel dimensions of [64, 128, 256, 512]. Each level of the encoder utilizes a series of residual blocks with 3×3 convolutional kernels and ReLU activation functions between the blocks.

Atrous Spatial Pyramid Pooling (ASPP) blocks with dilation rates of 2, 4, and 8 are used in each level of the spatial encoder to extract multi-scale features with the same spatial resolution as the feature maps extracted at the same level of the spatial encoder. ASPP blocks were used because they can extract contextual information from the image features at multiple scales of the original image. This information can then be used to accurately identify the objects of interest in the image to be fused:

$$F_{aspp} = Concat(F_{d=2}, F_{d=4}, F_{d=8}) \quad (1)$$

Attention Mechanism: We integrate squeeze-and-excitation (SE) attention blocks to adaptively recalibrate channel-wise feature responses. The attention weights are computed as:

$$s = \sigma\,(W_2 \cdot \delta(W_1 z)) \quad (2)$$

The decoder employs skip connections with attention gates that suppress irrelevant background regions while highlighting salient structures.

C - Frequency Domain Processing

The frequency branch operates on 16×16 overlapping blocks to maintain spatial adaptability while leveraging spectral energy compaction. For each block $B_{\{i,j\}}$ we apply 2D Discrete Cosine Transform:

$$F(u,v) = \frac{2}{\sqrt{MN}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} B(x,y)\, cos\ \left(\frac{\pi}{N}(x+\frac{1}{2})u\right) cos\ \left(\frac{\pi}{M}(y+\frac{1}{2})v\right) \quad (3)$$

$$\hat{F}(u,v) = \begin{cases} F(u,v), & if\ |\ F(u,v)\ | > \tau(u,v,\sigma) \\ 0, & otherwise \end{cases} \quad (4)$$

$$\tau(u,v,\sigma) = k\,\sigma\sqrt{2log\ (MN)} \quad (5)$$

The thresholding is performed differently on different frequency bands; lower thresholds are used on the low-frequency components and higher thresholds on the high-frequency components. After thresholding, the inverse DCT is used to reconstruct the frequency coefficients, which are then combined together using weighted averaging to eliminate blocking artifacts.

D - Edge-Aware Weight Estimation

The edge-aware module computes fusion weights based on local gradient magnitude to preserve structural boundaries. For each pixel location (x, y), the gradient magnitude is computed using Sobel operators as:

$$G(x,y) = \sqrt{G_x^2 + G_y^2} \quad (6)$$

where Gx and Gy represent the horizontal and vertical gradient components.

The spatial domain weight is computed using a sigmoid gating function:

$W_{s(x,y)} = \backslash frac\{1\}\{1\ +\ e^{\{-\alpha(G(x,y)-\beta)\}}\}$ (7)

where alpha controls the transition sharpness and beta represents the edge threshold (empirically set to 0.1 for normalized gradients).

High gradient regions (edges) yield $Ws \rightarrow 1$, favoring spatial domain preservation, while smooth regions yield $Ws \rightarrow 0$, favoring frequency domain noise suppression.

The frequency weight is computed as the complement:

$$G(x,y) = \sqrt{\{G_x^2 + G_y^2\}} \quad (8)$$

ensuring normalized fusion.

E - Self-Adaptive Fusion Strategy

The fusion module combines both domain outputs through learned residual refinement. The initial fused image is computed as:

$$I_{fuse}(x,y) = W_s(x,y)\, I_s(x,y) + W_f(x,y)\, I_f(x,y) \quad (9)$$

G - Theoretical and Stability Analysis

The proposed fusion strategy can be interpreted as a convex combination of spatial and frequency domain estimators. Specifically, the fused output is defined as a weighted sum of two complementary representations, where the weights are adaptively estimated based on local image structure. This formulation ensures that the fusion process remains bounded and stable, preventing amplification of noise artifacts .

Since the weights satisfy $0 \le W_s(x,y) \le 1$, the output lies within the convex hull of the input representations, guaranteeing numerical stability. Furthermore, by assigning higher weights to the domain with lower local reconstruction error (e.g., spatial near edges and frequency in smooth regions), the proposed method minimizes the expected denoising error while preserving structural details.

H - Algorithm Description

The complete SAFSF denoising procedure is summarized in Algorithm 1. First, the noisy image is processed in parallel through spatial and frequency-domain branches. The spatial branch extracts spatial features using a convolutional network. The frequency branch divides the image into blocks, applies a DCT transformation to each block, and applies a filter to each block's DCT coefficients.

The two feature maps are combined through an edge-aware weighting module, which calculates weights for each pixel according to the gradient magnitude at that pixel. These weights are used to combine the spatial and frequency-domain features. Finally, a residual network refines the resulting image to remove any remaining artifacts. The output of the SAFSF network is the denoised image.

Algorithm 1. Self-Adaptive Frequency–Spatial Fusion Denoising (SAFSF)

Input: Noisy image I_{noisy}, noise level σ(optional for blind denoising)

Output: Denoised image I_{clean}

1. Apply the spatial branch to the input image to obtain the spatial representation I_s.
2. Partition I_{noisy}into overlapping 16×16blocks $\{B_{i,j}\}$.
3. For each block $B_{i,j}$:
   a. Compute its 2D DCT coefficients $F_{i,j}$.
   b. Apply adaptive band-wise filtering to obtain $\hat{F}_{i,j}$.
   c. Reconstruct the filtered block $\hat{B}_{i,j}$using inverse DCT.
4. Aggregate all reconstructed blocks through overlap averaging to generate the frequency-domain output I_f.
5. Compute the gradient map $G$of the noisy image using the Sobel operator.
6. Estimate the spatial fusion weights $W_s$using sigmoid-based edge-aware gating.
7. Compute the complementary frequency weights $W_f = 1 - W_s$.
8. Fuse both domain outputs to generate the intermediate representation: $I_{fuse} = W_s \odot I_s + W_f \odot I_f$.
9. Apply the residual refinement module to obtain I_{res}.
10. Generate the final denoised output: $I_{clean} = I_{fuse} + I_{res}$.
11. Return I_{clean}.

4. Experimental Setup

A - Dataset Description

We evaluate the proposed framework on six benchmark datasets covering both synthetic and real-world scenarios:

Table 2. Dataset Specifications and Characteristics

Dataset	Images	Resolution	Type	Usage
Set12	12	256 × 256	Grayscale synthetic	Standard benchmark
BSD68	68	321 × 481	Grayscale synthetic	Validation
Kodak24	24	768 × 512	Color synthetic	Color denoising
Urban100	100	1024 × 1024	Color synthetic	Urban scenes
McMaster	18	500 × 500	Color synthetic	Textured images
SIDD	320 pairs	Variable	Real-world smartphone	Real noise evaluation

The SIDD dataset comprises real noisy images captured by five different smartphone cameras under various lighting conditions, providing 320 image pairs for training and 1280 validation patches (256×256) for evaluation.

B - Noise Models

We evaluate performance across multiple noise distributions:

Additive White Gaussian Noise (AWGN): Standard deviations $\sigma \in \{15,25,50\}$

Poisson Noise: Signal-dependent noise simulating low-light photography

Real-World Noise: SIDD dataset containing complex sensor noise patterns

For synthetic experiments, the training data is generated by adding noise to the clean images from the BSD400 and Waterloo Exploration Database (WED) containing 4744 images.

C - Evaluation Metrics

We employ three complementary metrics for comprehensive evaluation:

1. Peak Signal-to-Noise Ratio (PSNR):

$$PSNR = 10log_{10}\left(\frac{(2^B-1)^2}{MSE}\right) \quad (10)$$

2. Structural Similarity Index (SSIM):

$$c_1 = (k_1 L)^2, c_2 = (k_2 L)^2 \quad (11)$$

3. Learned Perceptual Image Patch Similarity (LPIPS): Deep learning-based perceptual metric using features from pre-trained VGG networks, better aligned with human visual perception than traditional metrics.

D - Implementation Details

The proposed framework is implemented in PyTorch and trained on NVIDIA A100 GPUs. Key implementation parameters:

Optimizer: Adam with initial learning rate 10^{-1}, decaying by 0.5 every 50 epochs

Batch Size: 16 for synthetic noise, 8 for real-world training

Patch Size: 128×128 for training, full resolution for testing

Epochs: 200 for synthetic, 300 for real-world fine-tuning

Data Augmentation: Random flips, rotations, and color jittering for color images

The DCT block size is set to 16×16 with 50% overlap (stride 8). Edge-aware parameters are fixed as $\alpha=10$ and $\beta=0.1$ across all experiments.

5. Results and Analysis

A - Quantitative Evaluation

Table 3. Performance Comparison

Method	BSD68 PSNR (dB)	BSD68 SSIM	Urban100 PSNR (dB)	Urban100 SSIM	Params (M)	FLOPs (G)
BM3D	27.36	0.763	27.93	0.840	-	-
DnCNN	27.92	0.789	27.59	0.833	0.668	43.9
FFDNet	27.98	0.790	28.05	0.840	0.852	14.0
BRDNet	28.16	0.794	28.56	0.858	1.2	50.0
Restormer	28.41	0.810	29.31	0.878	26.1	155.9
NAFNet	28.35	0.808	29.15	0.875	11.7	16.1
Pureformer	28.68	0.819	29.78	0.890	15.2	45.0
SAFSF (Ours)	28.85	0.825	30.05	0.895	18.5	35.2

Table 4. Performance on Color Benchmarks

Method	CBSD68 PSNR (dB)	Kodak24 PSNR (dB)	McMaster PSNR (dB)	Average SSIM
DnCNN	30.48	30.83	32.24	0.870
FFDNet	30.70	31.09	32.46	0.875
Restormer	34.39	31.78	28.61	0.935
NAFNet	34.35	31.75	28.58	0.933
SAFSF (Ours)	34.52	31.92	28.75	0.938

SAFSF consistently outperforms competing methods on color benchmarks.

Table 5. Real-World Performance

Method	SIDD PSNR (dB)	SIDD SSIM	DND PSNR (dB)
CBDNet	38.06	0.942	37.90
RIDNet	39.26	0.953	39.23
NAFNet	39.85	0.958	38.41

Restormer	40.03	0.959	38.65
Hybrid-CNN-T	40.05	0.961	39.91
SAFSF (Ours)	40.18	0.963	40.05

SAFSF achieves the best performance on real-world datasets (SIDD and DND).

B - Comparative Analysis

SAFSF achieves superior PSNR/SSIM performance compared to state-of-the-art methods while maintaining lower computational complexity than transformer-based models.

C - Statistical Validation

We perform statistical significance testing using paired t-tests and Wilcoxon signed-rank tests across all test images. Results indicate that SAFSF improvements over the second-best method (Pureformer) are statistically significant ($p<0.01$) for all noise levels $\sigma \geq 25$.

D - Ablation Study

Table 6. Ablation Study on BSD68 (σ=50)

Configuration	PSNR (dB)	SSIM	PSNR
Spatial Branch Only	27.95	0.792	-0.90
Frequency Branch Only	27.68	0.785	-1.17
Fixed Fusion (0.5 / 0.5)	28.42	0.812	-0.43
Without Edge-Aware	28.35	0.808	-0.50
Without Residual	28.62	0.818	-0.23
Full SAFSF (Proposed)	28.85	0.825	0.00

The ablation study confirms that: None of the two branches alone could achieve the best result, with the spatial branch outperforming the frequency branch by 0.27 dB. The Edge-aware weighting method brought an additional 0.50 dB of improvement to the spatial branch. The Residual learning method provided an additional 0.23 dB improvement on top of the Edge-aware weighting method. The Fixed 50/50 method resulted in the worst performance for the fusion method.

E - Computational Complexity Analysis

Table 7. Computational Efficiency Comparison

Method	Parameters (M)	FLOPs (G)	Runtime (ms)	PSNR/Param
DnCNN	0.668	43.9	45	41.8
FFDNet	0.852	14.0	28	32.8
NAFNet	11.7	16.1	35	2.4
Restormer	26.1	155.9	180	1.1
Pureformer	15.2	45.0	95	1.9
SAFSF (Ours)	18.5	35.2	72	1.6

SAFSF achieves favorable efficiency-performance trade-offs with the ability to process 512×512 images at 72ms on NVIDIA A100 GPUs. Additionally, PSNR per million parameters indicates that SAFSF has competitive parameter efficiency relative to other light-weight image super-resolution models.

6. Discussion

The effectiveness of SAFSF stems from its ability to adaptively combine the spatial and frequency domain representations based on the characteristics of the image areas. SAFSF achieves excellent results with different types of noise while offering a considerably lower computational complexity than transformer-based denoising methods. However, the introduction of artifacts at high levels of noise can be an issue for SAFSF, as the block-wise DCT implementation. Additionally, the estimation of the noise levels external to the algorithm may present issues in blind scenarios for the method.

7. Conclusion

A novel Self-Adaptive Frequency-Spatial Fusion (SAFSF) framework is presented in this paper for image denoising. By fusing spatial and frequency domain features with an edge-aware adaptive weighting scheme, the proposed framework achieves state-of-the-art denoising performance on both synthetic and real-world image collections. The framework surpasses existing methods by preserving edges while removing noise. Furthermore, SAFSF achieves this performance with competitive computational complexity. Thus, SAFSF establishes a new paradigm for image denoising methods that utilize both spatial and frequency domain representations.

References

[1] S. Akbar and A. Verma, “Analyzing Noise Models and Advanced Filtering Algorithms for Image Enhancement,” Oct. 29, 2024, Arxiv Org. doi: 10.48550/arxiv.2410.21946.

[2] R. Agrawal, “Analyzing the Performance of Image Denoising Techniques,” BJBio, vol. 13, no. 3, pp. 8–14, Nov. 2022, doi: 10.62865/bjbio.v13i3.51.

[3] M. Sharifymoghaddam, “Image Denoising in Spatial and Transform Domains,” May 2021, doi: 10.32920/ryerson.14665422.
[4] R. Paul, A. Sengupta, S. Purkait, and S. Khan, “Application of discrete domain wavelet filter for signal denoising,” De Gruyter, 2022, pp. 1–26. doi: 10.1515/9783110697216-001.
[5] L. Abdullaev, M. Tkachenko, and T. Nguyen, “Revisiting Transformers with Insights from Image Filtering and Boosting,” June 12, 2025, Arxiv Org. doi: 10.48550/arxiv.2506.10371.
[6] X. Tong, “A Study on Transformer Optimization for Image Processing on Edge Devices,” ACE, vol. 184, no. 1, pp. 7–15, Aug. 2025, doi: 10.54254/2755-2721/2025.ld25924.
[7] H. Pan et al., “Domain Generalization with Fourier Transform and Soft Thresholding,” Sept. 18, 2023, Arxiv Org. doi: 10.48550/arxiv.2309.09866.
[8] H. Pan et al., “Domain Generalization with fourier Transform and soft thresholding,” Institute Of Electrical Electronics Engineers, Apr. 2024, pp. 2106–2110. doi: 10.1109/icassp48485.2024.10446303.
[9] S. A. Elmasry, W. A. Awad, and S. A. Abd El-Hafeez, “Review of Different Image Fusion Techniques: Comparative Study,” Springer Singapore, 2020, pp. 41–51. doi: 10.1007/978-981-15-3075-3_3.
[10] A. Chaudhary, “Image Fusion Methods and Applications: A Review,” joit, vol. 2023, no. 1, Dec. 2023, doi: 10.61453/joit.v2023no14.
[11] B. Ponnan, R. S. Sreedivya, and G. Sreelatha, “A Glance On Various Image Denoising Techniques,” Institute Of Electrical Electronics Engineers, Aug. 2023, pp. 30–35. doi: 10.1109/icscc59169.2023.10335073.
[12] H. Li and D. Wu, “Hybrid-Domain Synergistic Transformer for Hyperspectral Image Denoising,” July 27, 2025, Cambridge University. doi: 10.48550/arxiv.2507.20099.
[13] H. Li and D. Wu, “Hybrid-Domain Synergistic Transformer for Hyperspectral Image Denoising,” Applied Sciences, vol. 15, no. 17, p. 9735, Sept. 2025, doi: 10.3390/app15179735.
[14] K. Xu, G. He, T. Qiao, and Z. Liu, “Adaptive Dual-Domain Debanding: A Novel Algorithm for Image and Video Enhancement,” Association for Computing Machinery, Oct. 2024, pp. 49–58. doi: 10.1145/3688863.3689572.
[15] X. Shen, Y. Zhang, H. Chen, and D. Gai, “Multi-focus noisy image fusion based on gradient regularized convolutional sparse representatione,” Association for Computing Machinery, Mar. 2021, pp. 1–7. doi: 10.1145/3444685.3446325.
[16] K. Hu, Q. Zhang, M. Yuan, and Y. Zhang, “SFDFusion: An Efficient Spatial-Frequency Domain Fusion Network for Infrared and Visible Image Fusion,” in Frontiers in artificial intelligence and applications, Institut F R Ost Und S Dosteuropaforschung, 2024. doi: 10.3233/faia240524.
[17] M. Sharifymoghaddam, “Image Denoising in Spatial and Transform Domains,” May 2021, doi: 10.32920/ryerson.14665422.v1.
[18] H. Hong, M. He, K. Wang, and L. Wu, “An Image Denoising Method for Real Scene Based on Pixel-Level Noise Estimation,” Association for Computing Machinery, Oct. 2022, pp. 306–311. doi: 10.1145/3569966.3570058.
[19] M. Snehalatha, D. N. Ramamurthy, K. Swetha, K. Vishnupriya, P. Sreelekha, and N. Niharika, “An Image Denoising in Spatial Domain using Bilateral Filter,” JOECS, vol. 7, no. 2, pp. 9–14, July 2022, doi: 10.46610/joecs.2022.v07i02.002.
[20] G. Eleyan, “Image denoising using deep learning: Comparative study,” 2576-8484, vol. 9, no. 7, pp. 1345–1359, July 2025, doi: 10.55214/25768484.v9i7.8918.
[21] C. Meng, H. Feng, Q. Li, and Z. Xu, “Spatial-Adaptive Network for Single Image Denoising,” Jan. 28, 2020, Arxiv Org. doi: 10.48550/arxiv.2001.10291.
[22] Z. Chen, P. Qin, J. Zeng, Q. Song, P. Zhao, and R. Chai, “LGIT: local–global interaction transformer for low-light image denoising,” Sci Rep, vol. 14, no. 1, p. 21760, Sept. 2024, doi: 10.1038/s41598-024-72912-z.
[23] S. Guo, H. Yong, X. Zhang, J. Ma, and L. Zhang, “Spatial-Frequency Attention for Image Denoising,” Feb. 27, 2023, Arxiv Org. doi: 10.48550/arxiv.2302.13598.
[24] I. Hashim Latif, S. Haider Abdulredha, and S. Khalid Abdul Hassan, “Discrete Wavelet Transform-Based Image Processing: A Review,” ANJS, vol. 27, no. 3, pp. 109–125, Sept. 2024, doi: 10.22401/anjs.27.3.13.
[25] U. Tuba and D. Zivkovic, “Image Denoising by Discrete Wavelet Transform with Edge Preservation,” Institute Of Electrical Electronics Engineers, July 2021. doi: 10.1109/ecai52376.2021.9515079.
[26] H. Q. Birdawod, A. M. Khudhur, D. H. Kadir, and D. M. Saleh, “A Wavelet Shrinkage Mixed with a Single-level 2D Discrete Wavelet Transform for Image Denoising,” KJAR, vol. 9, no. 2, pp. 1–12, July 2024, doi: 10.24017/science.2024.2.1.
[27] Z. Xiong, X. Zhang, Q. Hu, and H. Han, “IFormerFusion: Cross-Domain Frequency Information Learning for Infrared and Visible Image Fusion Based on the Inception Transformer,” Remote Sensing, vol. 15, no. 5, p. 1352, Feb. 2023, doi: 10.3390/rs15051352.
[28] Z. Wang, Z. Zhang, W. Qi, F. Yang, and J. Xu, “FreqGAN: Infrared and Visible Image Fusion via Unified Frequency Adversarial Learning,” IEEE Trans. Circuits Syst. Video Technol., vol. 35, no. 1, pp. 728–740, Jan. 2025, doi: 10.1109/tcsvt.2024.3460172.
[29] X. Luan et al., “FMambaIR: A Hybrid State-Space Model and Frequency Domain for Image Restoration,” IEEE Trans. Geosci. Remote Sensing, vol. 63, pp. 1–14, Jan. 2025, doi: 10.1109/tgrs.2025.3526927.
[30] J. Hu and Z. Wang, “A novel CNN architecture for image restoration with implicit frequency selection,” Connection Science, vol. 37, no. 1, Feb. 2025, doi: 10.1080/09540091.2025.2465448.
[31] B. Zheng et al., “Learning Frequency Domain Priors for Image Demoireing,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 11, pp. 7705–7717, Nov. 2022, doi: 10.1109/tpami.2021.3115139.
[32] Z. H. Shah et al., “Image restoration in frequency space using complex-valued CNNs.,” Front. Artif. Intell., vol. 7, p. 1353873, Sept. 2024, doi: 10.3389/frai.2024.1353873.