I.J. Image, Graphics and Signal Processing, **2026, 2, 68-86**
Published Online on April 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2026.02.05

Image Denoising in the Contourlet Domain Using RBF Network and Kalman Filter

Rachid Benoudi*
MSIA Team, IMIA Laboratory, Faculty of Sciences and Technology, Errachidia, Moulay Ismail University Meknes, Morocco
E-mail: r.benoudi@edu.umi.ac.ma
ORCID iD: https://orcid.org/0009-0007-2886-7270
*Corresponding Author

Youssef Qaraai
MSIA Team, IMIA Laboratory, Faculty of Sciences and Technology, Errachidia, Moulay Ismail University Meknes, Morocco
E-mail: y.qaraai@fste.umi.ac.ma
ORCID iD: https://orcid.org/0000-0003-0804-8167

Mohamed Ouhda
TIAD Laboratory, Departement of Computer and Mathematics, Higher School of Technology, khenifra, Morocco
E-mail: ouhda.med@gmail.com
ORCID iD: https://orcid.org/0000-0003-1615-367X

Received: 20 June, 2025; Revised: 18 September, 2025; Accepted: 14 January, 2026; Published: 08 April, 2026

Abstract: Image denoising remains a fundamental challenge in image processing, particularly when dealing with additive white gaussian noise (AWGN) that degrades visual quality and information content. This paper introduces a novel multi-stage denoising framework that uniquely combines Contourlet transform, radial basis function neural networks (RBFNN), and kalman filtering to effectively preserve important image features while removing noise. The contourlet transform first decomposes images into multi-resolution, directional subbands, providing a sparse representation that better captures geometric structures compared to traditional wavelet approaches. We then employ an RBFNN trained through back-propagation to adaptively threshold the contourlet coefficients based on local image characteristics and noise levels. Finally, kalman filtering is applied as a post-processing step to further suppress residual noise artifacts. Comprehensive experiments conducted on standard benchmark datasets demonstrate that our approach outperforms several state-of-the-art methods, including BM3D and recent deep learning-based techniques, particularly at moderate to high noise levels ($\sigma \geq 15$). Quantitative evaluations show our method achieves superior PSNR improvements of up to 2.4dB and SSIM improvements of 0.12 compared to recent competing approaches, while qualitative results confirm better preservation of edges and textural details. The proposed framework offers an effective balance between computational efficiency and denoising performance, making it suitable for various practical applications.

Index Terms: Image Restoration, Radial Basis Function Neural Networks, Contourlet Transform, Kalman Filters

1. Introduction

Digital image acquisition is inherently susceptible to various forms of degradation, significantly impacting image quality and subsequent analysis. Among these degradations, noise contamination remains a persistent challenge across numerous imaging applications, from medical diagnostics and remote sensing to consumer photography. The presence of noise not only diminishes visual quality but also hampers the effectiveness of subsequent image processing tasks such as segmentation, recognition, and feature extraction.

Image denoising—the process of estimating the underlying clean image from its noise-contaminated observation—represents a fundamental yet challenging inverse problem in image processing. Its inherent difficulty stems from the ill-posed nature of the problem: there exist numerous potential solutions that could have produced the observed noisy image. The challenge lies in recovering the clean image while preserving important structural elements such as edges,

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

textures, and fine details that carry critical information.

Among various noise models, additive white gaussian noise (AWGN) has received particular attention due to its mathematical tractability and prevalence in real-world scenarios. AWGN typically emerges from electronic circuit noise, sensor imperfections, or quantization errors during the analog-to-digital conversion process. The ubiquity of AWGN across imaging systems has made its removal a central focus in image restoration research.

Over the past several decades, numerous denoising approaches have emerged, each with distinct theoretical foundations and performance characteristics. Traditional spatial domain filters, such as Gaussian, mean, and median filters, while computationally efficient, often produce blurred images with compromised edge details. Frequency domain techniques like Fourier and wavelet-based methods offer improved performance through multi-scale analysis but may introduce artifacts around discontinuities due to their limited directional selectivity.

More recent advances have explored patch-based methods that exploit non-local similarities within images. Algorithms such as Non-Local Means (NLM) [1] and Block-Matching and 3D filtering (BM3D) [2] have demonstrated remarkable denoising capabilities by leveraging redundant patterns across an image. However, these approaches face significant computational demands and diminished effectiveness when handling images with limited self-similarities or unique textural patterns.

The emergence of deep learning has revolutionized image denoising, with Convolutional Neural Networks (CNNs) achieving impressive results through end-to-end training on large datasets. Models such as DnCNN [3], NLRN [4], and N3Net [5] have established new performance benchmarks by learning complex noise patterns directly from data. Nevertheless, these approaches require substantial computational resources for training, lack theoretical interpretability, and may struggle with noise levels or image characteristics not represented in their training data.

Despite these advances, significant challenges persist. Most notable is the fundamental trade-off between noise removal and detail preservation—aggressive denoising often leads to over-smoothing and loss of critical image features, while conservative approaches may leave residual noise. Additionally, many current methods perform inconsistently across different noise levels or image content types, limiting their practical utility.

In this paper, we propose a novel hybrid framework that addresses these limitations by integrating three complementary approaches: Contourlet transform, Radial basis function neural networks (RBFNN) and Kalman filtering. The Contourlet transform provides a sparse multi-resolution and multi-directional representation that effectively captures the intrinsic geometric structures in images, particularly curves and edges that conventional wavelet transforms struggle to represent efficiently [6]. Within this transform domain, we employ an adaptive RBFNN to discriminate between noise and significant image coefficients, leveraging the network's universal approximation capabilities [7] and efficient training characteristics. Finally, Kalman filtering refines the reconstructed image by addressing any residual noise artifacts, exploiting the temporal correlation and statistical properties of the noise.

Our key contributions include the development of a multi-stage denoising framework that synergistically combines the strengths of Contourlet Transform, RBFNN, and Kalman filtering to achieve superior noise removal while preserving critical image features. We introduce an adaptive RBFNN approach specifically designed for the Contourlet domain, with optimized training procedures and network architecture tailored for effective coefficient thresholding. Through comprehensive evaluation against state-of-the-art methods across various noise levels and image types, we demonstrate consistent performance improvements, particularly for moderate to high noise conditions where existing approaches typically struggle. Additionally, we provide an analysis of the individual contribution of each component in our pipeline through detailed ablation studies, offering insights into the synergistic effects of our multi-stage approach.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive review of related work in image denoising, highlighting the strengths and limitations of existing approaches. Section 3 presents the theoretical foundations of the Contourlet Transform and RBFNN, followed by a detailed description of our proposed methodology. Section 4 describes our experimental setup, including datasets, evaluation metrics, and implementation details, along with comprehensive results and comparisons with state-of-the-art methods. Finally, Section 6 concludes the paper with a summary of our findings and directions for future research.

2. Related Works

Image denoising has been extensively studied over the past several decades, resulting in a diverse array of methodologies. This section provides a comprehensive review of existing approaches, highlighting their strengths and limitations. We categorize the current image denoising algorithms into three broad groups: filter-based methods, model-based methods, and learning-based methods.

Filter-based denoising algorithms represent some of the earliest approaches to noise reduction and can be further subdivided into linear and non-linear filtering techniques. Linear filters, including mean filters [8, 9] and Wiener filters [10], offer computational simplicity and are mathematically tractable. The mean filter operates by replacing each pixel with the average value of its neighborhood, effectively smoothing the image. The Wiener filter attempts to minimize the mean square error between the estimated and original image by leveraging statistical properties of the image and noise. Despite their widespread application, linear filters suffer from significant drawbacks—they tend to blur important edge information and perform poorly when confronted with non-Gaussian noise distributions.

To address these limitations, researchers developed non-linear filtering techniques [11]. Among these, the filters proposed by Lee [12, 13] demonstrate superior edge-preservation capabilities by adapting their behavior based on local image statistics. These approaches better maintain structural information while removing noise, but still struggle with preserving fine textures and low-contrast details that are crucial for visual perception and subsequent analysis tasks.

Building upon the limitations of traditional filtering approaches, a significant advancement in image denoising came with the introduction of patch-based and non-local methods, which exploit the inherent redundancy and self-similarity present in natural images. Unlike pixel-based approaches, these methods operate on groups of similar patches. The NonLocal Means (NLM) algorithm [1] revolutionized denoising by estimating each pixel's true value through a weighted average of all pixels in the image, with weights determined by patch similarity. This approach significantly outperformed contemporary methods by preserving textures and repetitive structures.

Further refinement led to the development of Block-Matching and 3D filtering (BM3D) [2, 14], which combines patch similarity with transform-domain collaborative filtering. BM3D groups similar patches into 3D blocks, applies a 3D transform, performs shrinkage of the transform coefficients, and finally reconstructs the image through inverse transformation. Despite being introduced over a decade ago, BM3D remains a competitive baseline, approaching theoretical limits for natural images as demonstrated by [15] and [16].

While highly effective for images with recurring patterns, patch-based methods face challenges when processing unique or infrequently occurring patches [17, 18]. Additionally, their computational complexity increases substantially with image size, limiting applicability in time-sensitive applications.

Parallel to the development of patch-based methods, model-based approaches emerged that formulate image denoising as an optimization problem, typically attempting to minimize an energy function that incorporates both data fidelity and regularization terms. Among these, the Maximum a Posteriori (MAP) framework enables the integration of prior knowledge about image statistics. Total Variation (TV) regularization [19] assumes that clean images have minimal gradient magnitude, effectively preserving edges while removing noise. However, TV regularization tends to produce piecewise constant results, sometimes creating unnatural "staircase" artifacts in smooth regions.

Sparsity-based models [20, 21] exploit the observation that natural images have sparse representations in certain transform domains. By enforcing sparsity constraints, these approaches effectively separate noise from signal. Nevertheless, the performance of these methods heavily depends on the choice of dictionary or transform domain, which may not optimally represent all image structures.

More recently, the non-local self-similarity (NSS) prior [22, 23] has been incorporated into model-based frameworks, leveraging both local and non-local image statistics. While these approaches show improved performance, they often involve complex optimization problems with high computational demands and parameter sensitivity.

The evolution of denoising methods has been further transformed by the past decade's emergence of learning-based approaches, particularly those utilizing deep neural networks. These methods learn the mapping from noisy to clean images directly from data, circumventing the need for hand-crafted features or explicit modeling. Following the pioneering work of [24], numerous CNN-based methods have demonstrated state-of-the-art performance. DnCNN [3] employs residual learning and batch normalization to predict the noise component rather than the clean image, showing remarkable generalization across different noise levels. The Non-Local Recurrent Network (NLRN) [4] combines recurrent neural networks with non-local operations to capture both local and global image dependencies. N3Net [5] introduces a differentiable version of the nearest neighbor operation within a neural network framework, effectively integrating the strengths of non-local methods with deep learning.

Recent advances in deep learning-based denoising have focused on incorporating attention mechanisms and selfsupervised learning. Tian et al. [25] introduced attention-guided CNNs that adaptively focus on important image regions during denoising, improving detail preservation. Li et al. [26] proposed a dual mixture model based CNN that better captures the complex statistics of natural images and noise distributions. Neshatavar et al. [27] developed CVF-SID, a cyclic multi-variate function approach for self-supervised image denoising that disentangles noise from image content without requiring clean reference images.

Despite their impressive performance, learning-based approaches face several challenges. They typically require large amounts of training data, which may not be available for specialized imaging modalities. The "black-box" nature of deep neural networks limits theoretical understanding and interpretability. Furthermore, these methods may exhibit reduced effectiveness when confronted with noise characteristics or image types that deviate from their training distribution. Recent studies have also highlighted computational efficiency concerns, with many deep learning methods requiring substantial GPU resources and extended training times, limiting their practical deployment in resource-constrained environments.

Recognizing the complementary strengths of different denoising paradigms and the limitations of individual approaches, researchers have increasingly explored hybrid and multi-stage approaches. These methods sequentially or synergistically combine multiple techniques to overcome individual limitations. Transform domain methods utilizing wavelets have shown promise for capturing multi-scale information. However, standard wavelet transforms struggle with efficiently representing directional features such as curves and edges in natural images. To address this limitation, more sophisticated transforms such as the curvelet [28] and contourlet [6] transforms were developed to better capture geometric structures.

The contourlet transform, in particular, provides a sparse multi-resolution and multi-directional representation through a combination of Laplacian Pyramid decomposition and directional filter banks. While effective at representing

directional features, the raw contourlet coefficients remain susceptible to noise contamination, necessitating appropriate thresholding mechanisms.

Recent developments in hybrid approaches have incorporated deep learning components with traditional regularization techniques. Wu et al. [29] proposed a multi-level deep learning denoiser combined with tight frame regularization for medical image reconstruction, demonstrating improved performance over individual approaches. Chen et al. [30] developed Flex-DLD, a deep low-rank decomposition model with flexible priors for hyperspectral image denoising, effectively combining the advantages of low-rank modeling with deep learning adaptability.

A particularly notable development in hybrid approaches is the emergence of Deep Image Prior (DIP) methods. These approaches utilize untrained convolutional neural networks as implicit priors for image reconstruction, eliminating the need for large training datasets. However, recent studies have revealed significant limitations in DIP-based methods. Shi et al. [31] demonstrated that DIP suffers from spectral bias, tending to learn low-frequency components first, which can lead to artifacts with excessive iterations. The computational demands are substantial, with processing a single 512×512 image requiring approximately 2400 iterations on modern GPUs, making them impractical for real-time applications.

Furthermore, DIP methods face critical overfitting issues that require careful early stopping strategies, but determining optimal stopping points remains challenging without external validation data. Xu et al. [32] attempted to address some of these limitations by combining DIP with second-order total generalized variation, but the fundamental computational and stopping criteria challenges persist.

Recent developments in image denoising have increasingly explored transformer-based architectures and hybrid frameworks that integrate traditional priors with deep learning. Transformer-based models such as SwinIR, proposed by Liang et al. [33], employ Swin Transformer blocks to achieve strong performance across a range of image restoration tasks, while Restormer, introduced by Zamir et al. [34], leverages multi-head transposed self-attention to attain state-ofthe-art denoising results. Despite their effectiveness, these transformer-based approaches typically involve significantly higher computational complexity, often requiring five to ten times more parameters than conventional CNN-based methods, along with large-scale training datasets. In contrast, our approach prioritizes interpretability and efficiency by explicitly operating in the transform domain and relying on lightweight learning components. Alongside transformer-based methods, hybrid and self-supervised approaches have gained attention for their ability to combine data-driven learning with model-based regularization. Wu et al. [35] demonstrated the effectiveness of integrating multi-level deep learning denoisers with tight frame regularization for medical image reconstruction, highlighting the benefits of combining complementary priors. Similarly, Chen et al. [36] proposed Flex-DLD, which incorporates flexible priors within a deep low-rank decomposition framework for hyperspectral image denoising. These studies support the growing consensus that hybrid strategies can offer improved robustness and generalization. While sharing this philosophy, our method distinguishes itself by emphasizing classical transform-domain representations and shallow neural networks, enabling a more transparent and interpretable denoising process without the heavy computational burden of large-scale deep architectures. Neural networks have been employed for adaptive processing in transform domains, with radial basis function neural networks (RBFNNs) demonstrating particular promise due to their universal approximation properties [7, 37, 38]. When combined with appropriate post-processing techniques such as Kalman filtering, these hybrid approaches can potentially overcome the limitations of individual methods while maintaining computational efficiency.

3. Methodology

This section presents a detailed description of our proposed multi-stage image denoising framework. We begin with a formal problem formulation, followed by an in-depth explanation of each component in our approach: the Contourlet Transform, the Radial Basis Function Neural Network (RBFNN) for adaptive thresholding, and the Kalman filtering for post-processing refinement. We also discuss the integration of these components and the algorithmic implementation.

3.1. Motivation

Despite the significant advances in image denoising, several challenges persist, particularly with recent state-of-theart methods. The fundamental trade-off between noise suppression and detail preservation remains difficult to optimize, particularly for images with complex textures and fine structures, and recent attention-guided methods [25] attempt to address this but still struggle with maintaining consistent performance across varying image content types. Many existing methods exhibit inconsistent performance across different noise levels and image content types, and while deep learning approaches excel at low to moderate noise levels, their performance deteriorates significantly at higher noise intensities, limiting their practical applicability in challenging scenarios. Recent Deep Image Prior methods [31, 32], while promising for unsupervised learning, suffer from critical computational efficiency issues where processing a single 512×512 image requires approximately 2400 GPU iterations, making them impractical for real-time applications. Additionally, these methods face severe overfitting problems that necessitate careful early stopping strategies, but determining optimal stopping points remains challenging without validation data. Deep learning approaches [27, 26], while achieving impressive results, often lack interpretability and may require retraining for

different noise conditions, and the "black-box" nature of these methods makes it difficult to predict their behavior in novel scenarios or to incorporate domain-specific knowledge effectively. Computationally efficient methods typically sacrifice denoising quality, while high-performance approaches often demand substantial computational resources, and recent hybrid methods [29, 30] attempt to balance this trade-off but often introduce additional complexity in parameter tuning and model selection. Current hybrid approaches that combine traditional and learning-based methods often struggle with optimal integration of different components, leading to suboptimal performance or increased computational overhead.

Our work addresses these limitations by proposing a novel hybrid framework that combines the directional sensitivity of the contourlet transform, the adaptive learning capabilities of RBFNNs, and the statistical refinement of Kalman filtering. Unlike computationally intensive DIP methods that require thousands of iterations per image, our RBFNN-based adaptive thresholding provides efficient learning with clear stopping criteria. In contrast to "black-box" deep learning approaches, each component of our framework has clear theoretical justification and interpretable functionality. Our method demonstrates superior performance at moderate to high noise levels ($\sigma \geq 15$) where recent deep learning methods typically struggle, while maintaining computational efficiency suitable for practical applications. This approach aims to achieve superior denoising performance while maintaining computational efficiency and theoretical interpretability.

3.2. Theoretical Background

This section presents the theoretical foundations of the three complementary components integrated in our approach: Contourlet Transform for sparse representation, RBFNN for adaptive thresholding, and Kalman filtering for post-processing refinement.

Contourlet Transform Decomposition: The Contourlet Transform (CT) provides sparse multi-resolution and multidirectional representation, offering advantages over traditional wavelets in capturing directional information. The CT combines the Laplacian Pyramid (LP) for multi-scale decomposition with the Directional Filter Bank (DFB) for directional analysis. The LP decomposition by Burt and Adelson [39] filters and downsamples the original image f, while the DFB by Bamberger and Smith [40] decomposes high-frequency components into 2^l directional subbands. The sampling matrices for subbands at level l are:

$$S_k^{(l)} = \begin{cases} \mathrm{diag}(2^{l-1}, 2) & , 0 \leq k < 2^{l-1} \\ \mathrm{diag}(2, 2^{l-1}) & , 2^{l-1} \leq k < 2^l \end{cases} \quad (1)$$

RBFNN-Based Adaptive Thresholding: To address varying characteristics across image regions and noise levels, we employ Radial Basis Function Neural Networks for adaptive coefficient thresholding. The RBFNN utilizes universal approximation capabilities [7] with a three-layer architecture: input layer (contourlet coefficients), hidden layer (RBF units), and output layer (denoised estimates). Each hidden neuron employs a Gaussian activation function:

$$\phi_i(\|x - c_i\|) = \exp\left(-\frac{\|x-c_i\|^2}{2\sigma_i^2}\right) \quad (2)$$

where x represents input coefficients, c_i is the i-th RBF center, and σ_i is the width parameter. The network output is:

$$y_k = \sum_{j=1}^{n} \omega_{jk}\, \phi_j(x) \quad (3)$$

where ω_{jk} connects the j-th hidden neuron to the k-th output neuron. We train separate RBFNNs for different subbands using noisy-clean coefficient pairs, employing two-phase training: clustering for RBF centers and supervised learning for weights.This section provides comprehensive details on our RBFNN implementation to ensure reproducibility. For each directional subband s at decomposition level l with direction k, we train a dedicated RBFNN $N_{l,k}$ with the following architecture:

- **Input Layer:** Each input consists of a spatial neighborhood of contourlet coefficients. Specifically, for coefficient $c_{i,j}$ at position (i, j), we extract a $P \times P$ patch centered at (i, j), which is then vectorized into an input vector $\mathrm{x} \in \mathbb{R}^{P^2}$. We use $P = 8$ based on cross-validation.
- **Hidden Layer:** The hidden layer contains N_h radial basis function units, where N_h is adaptively determined during training (typically 50-150 neurons depending on subband complexity). Each hidden neuron j implements a Gaussian RBF:

$$\phi_j(\mathrm{x}) = \exp\left(-\frac{\|x-c_j\|^2}{2\sigma_j^2}\right) \quad (4)$$

where $c_j \in \mathbb{R}^{P^2}$ is the center of the j-th RBF and σ is its spread parameter.
- **Output Layer:** Single output neuron producing the denoised coefficient estimate:

$$\hat{c}_{i,j} = \sum_{j=1}^{N_h} w_j \, \phi_j(\mathrm{x}) + b \tag{5}$$

where w_j are the output weights and b is the bias term.

To ensure the network generalizes across various textures and noise intensities, training data is prepared as follows:

1. **Image Selection:** 200 clean images from BSD500 training set
2. **Noise Injection:** Each clean image u is corrupted with AWGN at multiple noise levels $\sigma \in \{5, 10, 15, 20, 25, 30\}$ to create noisy version $f = u + N(0, \sigma^2)$
3. **Contourlet Decomposition:** Both clean and noisy images undergo contourlet decomposition with structure [2, 4, 4, 4]
4. **Patch Extraction:** For each subband (l, k):
- Extract overlapping 8 ×8 patches with stride 4
- From noisy coefficients: input patches $\{x_n\}_{n=1}^{N}$
- From clean coefficients: target values $\{y_n\}_{n=1}^{N}$
- This yields approximately 50,000 training pairs per subband
5. **Normalization:** All patches are normalized to zero mean and unit variance before training we employ a two-phase training strategy to optimize the network parameters efficiently.

Phase 1: RBF Center Initialization via K-Means

The RBF centers are initialized using K-means clustering on the noisy input patches:

Algorithm 1. RBF Center Initialization

Require: Noisy patches $\{x_n\}_{n=1}^{N}$, target number of centers $K_{\max} = 150$

1: Randomly sample 10,000 patches from $\{x_n\}$ for efficiency

2: Apply K-means clustering with $K = K_{\max}$

3: Set initial centers: $\mathbf{c}_j = \mu_j$ (cluster centroids), $j = 1, \ldots, K$

4: Initialize spread parameters: $\sigma_j = \frac{d_{max}}{\sqrt{2K}}$

5: where $d_{\max}$ is maximum distance between any two centers **return** $\{c_j, \sigma_j\}_{j=1}^{K}$

Phase 2: Supervised Weight Optimization

With fixed centers and spread parameters, we optimize output weights using orthogonal least squares (OLS) regression:

$$\mathrm{w}^{*} = \arg\min_{\mathrm{w}} \sum_{\mathrm{n}=1}^{\mathrm{N}} \left(\mathrm{y_n} - \sum_{\mathrm{j}=1}^{\mathrm{N_h}} \mathrm{w_j} \, \phi_{\mathrm{j}}(\mathrm{x_n}) \right)^2 + \lambda \|\mathrm{w}\|^2 \tag{6}$$

where $\lambda = 0.001$ is the regularization parameter to prevent overfitting.

Adaptive Neuron Selection: We use an incremental approach:

1. Start with $N_h = 10$ neurons
2. Compute training MSE
3. If MSE > goal (0.01) and $N_h < 150$:
- Add neurons corresponding to largest error clusters
- Retrain weights
- Repeat until goal met or max neurons reached

The spread parameter σ critically affects performance. We select it via 5-fold cross-validation on the BSD500 validation set:

- Test values: $\sigma \in \{0.1, 0.5, 1.0, 2.0, 5.0\}$
- Metric: Average PSNR on validation set
- Selected value: σ = 1.0 (provides best bias-variance trade-off)

The implementation was carried out in MATLAB R2017a using the Neural Network Toolbox (specifically a modified newrb function). Training was performed on an AMD Ryzen 5 5600GE CPU with 8GB RAM, requiring approximately 5 minutes per subband. In total, 52 separate RBFNN models were trained to account for the 4 decomposition levels and their respective directional subbands.

Kalman Filter Post-Processing: After RBFNN processing and inverse CT reconstruction, Kalman filtering provides statistical refinement. We model image restoration as a state estimation problem where the clean image represents the true state. The state update equation is:

with Kalman gain matrix:

$$I_t = I_{t-1} + K_t(Y_t - HI_{t-1}) \tag{7}$$

$$K_t = P_{t|t-1}H^T(HP_{t|t-1}H^T + R)^{-1} \tag{8}$$

where $P_{t|t-1}$ is the predicted error covariance and R is the observation noise covariance. This stage exploits temporal correlation of residual noise, adaptively adjusts filtering strength, and preserves edges through statistical prior information.

3.3. Problem Formulation

Let us denote an original clean image as $u \in \mathbb{R}^{M\times N}$, where M and N represent the image dimensions. The observed noisy image $f \in \mathbb{R}^{M\times N}$ is modeled as:

$$f = u + b \tag{9}$$

where $b \in \mathbb{R}^{M\times N}$ represents Additive white gaussian noise (AWGN) with zero mean and standard deviation σ. When σ is unknown, as is common in practical scenarios, it can be estimated using the Fast Noise Variance Estimation method [41].

The objective of image denoising is to estimate the original clean image $\hat{u}$ from the noisy observation f, such that $\hat{u}$ approximates u as closely as possible. This estimation problem is ill-posed because multiple potential solutions could have generated the same noisy observation. Therefore, appropriate regularization is necessary to constrain the solution space.

Our approach addresses this challenge through a three-stage process:

1. Transformation of the noisy image into a sparse multi-resolution, multi-directional representation using the Contourlet Transform
2. Adaptive thresholding of the contourlet coefficients using a trained RBFNN
3. Refinement of the reconstructed image using Kalman filtering to further suppress residual noise

3.4. Proposed Method

In this section, we describe the image denoising approach proposed in this study. Let u denote the original image and b represent additive white Gaussian noise with zero mean and a standard deviation σ. In the absence of a known σ, it can be estimated using the Fast Noise Variance Estimation method as described in [41]. The noisy image is modeled as shown in 9.

Our objective is to develop an algorithm that effectively removes the noise from f and recovers a denoised image $\hat{u}$ that closely approximates u. The proposed denoising process consists of three sequential stages.

First, the contourlet transform is applied to f, decomposing it into multiscale directional coefficients that capture the geometric structure of the image and the information about the edges. This transformation yields a sparse representation that is more amenable to noise separation than the original spatial domain.

Next, denoising is performed in the contourlet domain using a Radial Basis Function Neural Network (RBFNN). Let c denote the contourlet coefficients of f. The RBFNN is trained on contourlet coefficients extracted from a set of images with known noise levels, thereby learning an adaptive thresholding function $T(\cdot)$. This function effectively distinguishes between the noise and signal components, resulting in denoised coefficients given by $c_{denoised} = T(c)$.

Subsequently, the inverse contourlet transform is used to reconstruct the denoised image $\hat{u}$ from the denoised coefficients. Since this reconstruction might still exhibit minor artifacts or residual noise, a Kalman filter is applied to $\hat{u}$. The Kalman filter operates recursively, refining the estimate of $\hat{u}$ by correcting for any discrepancies introduced during the inverse transformation. This final step ensures that the noise is effectively suppressed while preserving critical structural and edge information.

Through the combination of the contourlet transform, RBFNN-based adaptive thresholding in the transform domain, and subsequent Kalman filtering, the method produces a high-quality denoised image that approximates the original u with improved fidelity. Our complete denoising framework integrates the three components described above into a cohesive pipeline, as illustrated in Figure 1 and formally defined in Algorithm 2.

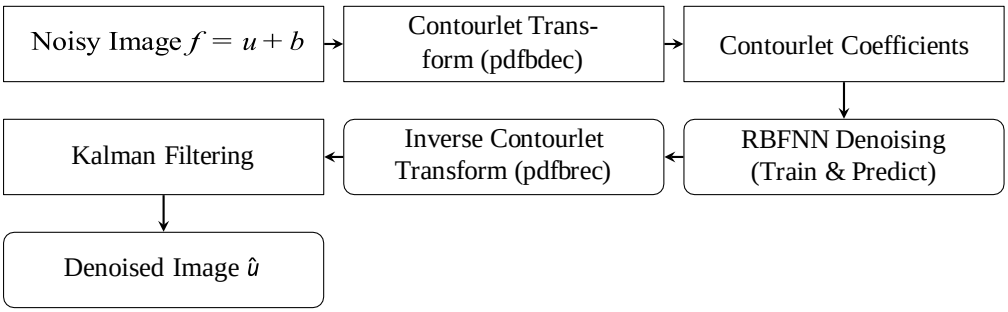


Fig. 1. Schema of the proposed multi-stage denoising framework integrating Contourlet Transform, RBFNN adaptive thresholding, and Kalman filtering.

Algorithm 2. Multi-Stage Image Denoising using Contourlet Transform, RBFNN, and Kalman Filtering

Require: Noisy image f, noise standard deviation σ (or estimate using [41])
Ensure: Denoised image $\hat{u}$
Stage 1: Contourlet Transform Decomposition
Decompose f using the Contourlet Transform:

$$C_{\text{noisy}} \leftarrow \text{pdfbdec}(f, \text{pfilt}, \text{dfilt}, \text{nlevels})$$

Stage 2: RBFNN-based Adaptive Thresholding
for each directional subband $i = 1, \ldots, N_{\text{subbands}}$ **do**
Extract noisy coefficients: $X_{\text{noisy}} \leftarrow C_{\text{noisy},i}$
Load pre-trained RBFNN model N_i for subband i
Predict denoised coefficients:

$$X_{\text{denoised}} \leftarrow \text{N}i(X_{\text{noisy}})$$

Update: $C_{\text{denoised},i} \leftarrow X_{\text{denoised}}$
end for
Stage 3: Image Reconstruction
Reconstruct the intermediate denoised image:

$$\hat{u}_{\text{intermediate}} \leftarrow \text{pdfbrec}(C_{\text{denoised}}, \text{pfilt}, \text{dfilt})$$

Stage 4: Kalman Filtering Refinement
Initialize Kalman filter parameters based on estimated residual noise
Apply Kalman filtering to $\hat{u}_{\text{intermediate}}$ for N_{iter} iterations to obtain $\hat{u}$
return $\hat{u}$

The key aspects that differentiate our approach from existing methods include our utilization of multi-resolution and multi-directional representation through the Contourlet Transform, which provides a more efficient representation of directional features compared to traditional wavelets, allowing for better separation of signal and noise. Our method employs adaptive localized thresholding where, instead of applying a global threshold, our RBFNN-based approach adapts to the local characteristics of different image regions and directional subbands. We incorporate statistical refinement through the Kalman filtering stage that leverages the statistical properties of the residual noise to further enhance the denoising performance without compromising edge and texture preservation. Most importantly, our approach achieves complementary strengths integration by combining transform-domain sparsity, neural network adaptability, and statistical filtering, thereby addressing the limitations of individual methods through their synergistic combination.

4. Simulation Results and Discussion

In this section, we present a comprehensive evaluation of our proposed multi-stage denoising framework. First, we describe the experimental setup, including datasets, parameter configurations, and evaluation metrics. We then analyze the performance of our method across different noise levels and image characteristics. Finally, we conduct a detailed comparison with state-of-the-art denoising techniques to demonstrate the effectiveness of our approach.

4.1. Experimental Setup

To ensure a thorough and unbiased evaluation, we conducted experiments on multiple standard benchmark datasets. We used a set of widely recognized test images including Barbara, Boat, Cameraman, and Man. To expand beyond the limited standard test set, we incorporated images from the Berkeley Segmentation Dataset (BSD500), which was used both for training the RBFNN models and for additional testing. For more comprehensive evaluation, we also tested our method on the Set14 and Urban100 datasets.

For the Contourlet Transform implementation, we utilize the framework developed by Do and Vetterli [6]. The decomposition structure employs a pyramid level configuration of [2, 4, 4, 4], which provides an effective balance between directional selectivity and computational efficiency. The ”pkva” filter is selected for the Laplacian pyramid stage, while the ”9-7” filter is used for the directional filter bank stage.

The RBFNN is implemented using MATLAB’s Neural Network Toolbox, specifically the ”newrb” function which creates a radial basis network with dynamic hidden layer sizing. For training, we employ a diverse set of clean reference images, artificially corrupted with AWGN at various noise levels (σ = 5, 10, 15, 20, 25, and 30). The network’s parameters are optimized using a combination of unsupervised clustering for center determination and supervised learning for weight optimization.

For the Kalman filtering stage, we estimate the residual noise characteristics after the RBFNN processing using the method described in [41]. The filter parameters, including the process and observation noise covariances, are adaptively tuned based on these estimates to achieve optimal performance.

The entire denoising framework is implemented in MATLAB, with careful optimization to maintain reasonable computational efficiency while preserving the denoising performance.

Having established our evaluation datasets, we implemented a systematic training protocol for the neural network component. For the training phase of the RBFNN, we used 200 clean image patches randomly selected from the BSD training set. These patches were artificially corrupted with AWGN at various noise levels to create noisy-clean image pairs for supervised learning.

To comprehensively assess the robustness and effectiveness of our approach across varying conditions, we designed experiments with multiple noise intensities. To evaluate the robustness of our method across different noise conditions, we artificially corrupted the clean images with AWGN at six different standard deviation levels: $\sigma = 5$, 10, 15, 20, 25, and 30. This range covers low to high noise conditions commonly encountered in practical applications.

With the datasets and noise conditions established, we proceeded to configure the algorithmic parameters for optimal performance. The key parameters of our method were configured as shown in Table 1.

Table 1. Parameter configuration for the proposed method.

Component	Parameter	Value/Configuration
Contourlet Transform	Decomposition levels	[2, 4, 4, 4]
	LP filter	9-7 filter
	DFB filter	pkva filter
RBFNN	Function	newrb (MATLAB)
	Spread parameter	1.0
	Goal MSE	0.01
	Maximum neurons	100 (dynamically adjusted)
Kalman Filter	Process noise covariance	0.01
	Observation noise covariance	Estimated from residual noise
	Iterations	5

The rationale behind these parameter choices is based on extensive preliminary experimentation and theoretical considerations. The decomposition structure with levels [2, 4, 4, 4] provides an effective multi-scale and multi-directional representation of the images. The combination of ”9-7” filter for the Laplacian pyramid stage and ”pkva” filter for the directional filter bank stage was chosen based on preliminary experiments that demonstrated its ability to efficiently capture directional features while maintaining computational efficiency. The RBFNN configuration utilizes dynamic neuron adjustment during training based on the error goal, while the Kalman filter observation noise covariance is estimated from the residual noise after the RBFNN stage.

Finally, to ensure comprehensive and objective assessment of denoising performance, we employed multiple evaluation criteria. To quantitatively assess the denoising performance, we employed the most widely used metrics Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mean Squared Error (MSE). In addition to these quantitative metrics, we also conducted visual inspections to assess the subjective quality of the denoised images, particularly focusing on the preservation of edges, textures, and fine details.

4.2. Performance Analysis of the Proposed Method

We first evaluate the denoising performance of our proposed method across various noise levels and image types. Figure 2 illustrates the visual quality of our denoising results on three representative test images (Cameraman, Barbara, and Boat) at different noise levels ($\sigma = 10$, 20, 30).

As observed from Figure 2, our method effectively removes noise while preserving important image features. Even at high noise levels ($\sigma = 30$), structural integrity and edge details are well maintained. For the Cameraman image, the fine details of the photographer, camera equipment, and tripod structure remain clearly visible with sharp edge definition. In the Barbara image, the complex textile patterns are preserved despite their high-frequency nature, which is particularly challenging for many denoising methods. The Boat image demonstrates our method’s ability to retain fine details such as ropes and mast structures even under severe noise conditions.



Fig. 2. Visual results of our proposed denoising method on standard test images. For each test case: the leftmost column shows the original image; the top row displays noisy versions corrupted by Gaussian noise at different noise levels ($\sigma = 10, 20, 30$); and the bottom row presents the corresponding denoised results obtained using our method. The results demonstrate the algorithm's effectiveness in preserving important image features while removing noise across various noise intensities and image textures.

The quantitative performance of our method across all test images and noise levels is presented in Table 2. The table reports PSNR (dB), SSIM and MSE values for each configuration.

Table 2. Quantitative performance of our proposed method on standard test images across various noise levels.

Metrics	$\sigma = 5$	$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$
Barbara						
PSNR	25.98	25.79	25.51	24.10	24.56	23.90
SSIM	0.8107	0.7856	0.7498	0.7010	0.6464	0.5870
MSE	0.0025	0.0026	0.0028	0.0031	0.0035	0.0040
Boat						
PSNR	32.17	31.52	30.56	29.38	28.10	26.78
SSIM	0.9032	0.8693	0.8209	0.7574	0.6858	0.6141
MSE	0.0006	0.0007	0.0009	0.001	0.001	0.0021
Cameraman						
PSNR	30.92	30.38	29.64	28.73	27.67	26.60
SSIM	0.9120	0.8736	0.8182	0.7501	0.6771	0.6013
MSE	0.0008	0.0009	0.0011	0.0013	0.0017	0.0021
Man						
PSNR	31.13	30.43	29.49	28.34	27.09	25.87
SSIM	0.8936	0.8646	0.8220	0.7656	0.6995	0.6329
MSE	0.0007	0.0009	0.0011	0.0015	0.0020	0.0026

From Table 2, we observe that our method consistently achieves high PSNR and SSIM values across all test images and noise levels. As expected, the performance metrics gradually decrease with increasing noise intensity. However, the rate of degradation is relatively modest, indicating the robustness of our approach even under severe noise conditions.

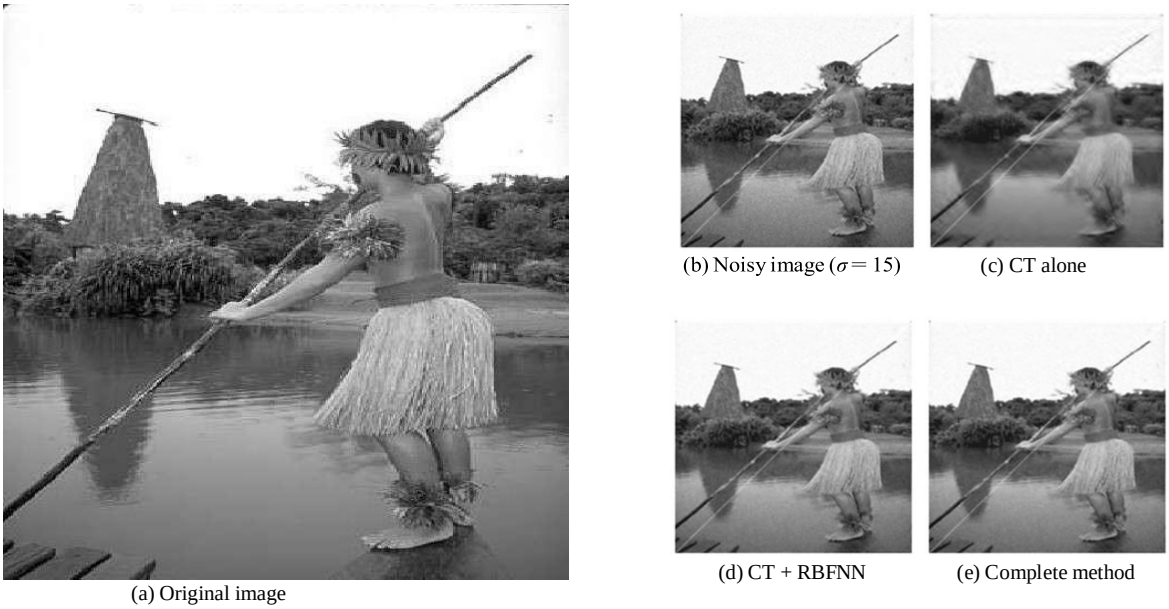
Notably, the method performs particularly well on images with mixed content types (e.g. Man), achieving PSNR values above 24 dB even at high noise levels ($\sigma = 30$). For images with complex textures (e.g., Barbara), the performance is slightly lower but still competitive, demonstrating our method's ability to handle challenging image structures.

4.3. Ablation Study

To demonstrate the contribution of each component in our multi-stage framework, we conducted an ablation study by evaluating:

1. Contourlet transform with basic thresholding alone (CT)
2. Contourlet transform with RBFNN thresholding but without Kalman filtering (CT+RBFNN)
3. Our complete method with all components (CT+RBFNN+Kalman)

Figure 3 shows the visual comparison of these configurations on the Man image corrupted with noise at $\sigma = 15$, while Table 3 provides the corresponding quantitative results.



(a) Original image (b) Noisy image ($\sigma = 15$) (c) CT alone (d) CT + RBFNN (e) Complete method

Fig. 3. Ablation study showing the contribution of each component in our proposed framework. (a) Original image; (b) Noisy image ($\sigma = 15$); (c) Contourlet transform with basic thresholding; (d) Contourlet transform with RBFNN adaptive thresholding; (e) Complete method with Kalman filtering refinement.

Table 3. Quantitative results of the ablation study on the Man image with σ = 15.

Method	PSNR (dB)	SSIM	MSE
Contourlet transform alone	25.98	0.7376	0.0025
Contourlet + RBFNN	29.19	0.8019	0.0014
Complete method (CT + RBFNN + Kalman)	29.49	0.8220	0.0011

The ablation study results clearly demonstrate the incremental improvement contributed by each component of our framework. The basic Contourlet transform with simple thresholding achieves reasonable denoising performance (PSNR = 25.98 dB) but suffers from residual noise and some loss of detail. The integration of RBFNN for adaptive thresholding significantly improves the results (PSNR = 29.19 dB, SSIM = 0.8019), with better edge preservation and reduced artifacts. Finally, the complete method with Kalman filtering further refines the result (PSNR = 29.49 dB, SSIM = 0.8220), particularly in removing residual noise while preserving fine details.

Visually, the progression from basic CT to the complete method shows noticeable improvements in edge definition, texture preservation, and overall image clarity. The RBFNN contribution is particularly significant, as evidenced by the substantial improvement in both PSNR and SSIM metrics compared to the basic CT. The Kalman filtering stage provides a more subtle but still valuable refinement, especially in homogeneous regions where residual noise might otherwise persist.

To rigorously assess each component's contribution, we expanded our ablation study across multiple images and noise levels.

Based on the extended ablation results in Table 4, the performance gains of the proposed pipeline can be clearly attributed to the complementary roles of the RBFNN and Kalman filtering stages. Across all test images and noise levels, the RBFNN acts as the primary performance driver, yielding substantial PSNR improvements over the CT-only baseline, with average gains of approximately 2.4 dB, 3.4 dB, and 2.7 dB at noise levels $\sigma = 10$, 20, and 30, respectively, alongside notable SSIM increases. These gains become more pronounced at higher noise levels, indicating a strong noise-dependent learning capability of the RBFNN. In contrast, the Kalman filter consistently provides a smaller but stable refinement, contributing an additional 0.24–0.50 dB in PSNR and modest SSIM improvements across all conditions. Importantly, the full CT + RBFNN + Kalman pipeline consistently outperforms either component alone, suggesting a synergistic effect in which the Kalman filter effectively refines the RBFNN output, leading to overall performance gains that exceed the sum of individual contributions.

Table 4. Extended ablation study across images and noise levels.

Method	PSNR (dB)			SSIM		
	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$
McMaster 1						
CT only	24.68	22.29	20.99	0.6972	0.5869	0.5251
CT + RBFNN	28.18	26.89	25.15	0.8422	0.7638	0.6725
CT + RBFNN + Kalman	**28.38**	**27.07**	**25.26**	**0.8462**	**0.7778**	**0.6812**
McMaster 17						
CT only	26.82	24.61	23.54	0.6919	0.5754	0.5176
CT + RBFNN	31.12	28.89	25.95	0.8521	0.7615	0.6238
CT + RBFNN + Kalman	**31.52**	**29.37**	**26.95**	**0.8677**	**0.7742**	**0.6539**
McMaster 18						
CT only	26.17	23.52	22.01	0.7019	0.6014	0.5466
CT + RBFNN	25.49	24.84	23.56	0.8035	0.7166	0.6207
CT + RBFNN + Kalman	**25.61**	**25.04**	**23.96**	**0.8135**	**0.7366**	**0.6307**
Average Improvement						
RBFNN contribution	+2.36	+3.4	+2.71	+0.1356	+0.1594	+0.1092
Kalman contribution	+0.24	+0.28	+0.5	+0.009	+0.015	+0.016

4.4. Extended Dataset Results

To validate the robustness and generalizability of our method beyond the standard test images, we evaluated its performance on the extended datasets mentioned in Section 4.1. Table 5 presents the average PSNR and SSIM values across different image categories and noise levels.

Based on the results in Table 4, the proposed method demonstrates consistent and robust performance across all extended datasets and noise levels. The McMaster and Urban100 datasets achieve the highest PSNR and SSIM values at lower noise levels, indicating strong denoising effectiveness on images with rich textures and structured content. As the noise level increases, a gradual degradation in performance is observed across all datasets, which is expected; however, the relative ranking remains stable. Urban100, characterized by complex urban and architectural patterns, exhibits slightly lower PSNR and higher MSE at higher noise levels compared to McMaster and BSD500, reflecting the

increased difficulty of preserving fine geometric details under strong noise. Despite these challenges, the proposed method maintains competitive PSNR, SSIM, and MSE values across all datasets, including Set14 and BSD500, underscoring its strong generalization capability to diverse image contents and structural complexities.

Table 5. Performance of our method on extended datasets (average values).

Metrics	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$
McMaster			
PSNR	31.20	29.04	26.52
SSIM	0.8655	0.7576	0.6203
MSE	0.0009	0.0013	0.0023
Urban 100			
PSNR	31.16	28.88	26.32
SSIM	0.8917	0.8014	0.6814
MSE	0.0011	0.0016	0.0026
Set14			
PSNR	29.80	28.08	25.89
SSIM	0.8451	0.7525	0.6268
MSE	0.0012	0.0017	00.27
BSD500			
PSNR	30.86	28.82	26.34
SSIM	0.8626	0.7606	0.6238
MSE	0.0009	0.0014	0.0023

5. Comparative Analysis

To position our work within the current landscape of image denoising research, we conducted comprehensive comparisons with several state-of-the-art methods. The selected comparison methods represent different denoising paradigms and include BM3D [2], a widely recognized benchmark method that employs collaborative filtering in the transform domain and, despite its age, remains competitive and serves as a standard reference point; DnCNN [3], a deep learning-based approach that uses residual learning and batch normalization to predict the noise component; and NLRN [4], a non-local recurrent network that combines recurrent neural networks with non-local operations.

Having established our comparative framework, we implemented rigorous experimental protocols to ensure reliable and unbiased evaluation. To ensure fair comparisons, all methods were evaluated using identical test images, noise conditions, and evaluation metrics. For deep learning-based methods, we used the pre-trained models provided by the respective authors, which were trained on the same noise levels as used in our experiments.

With the experimental setup established, we proceed to present the quantitative evaluation results. Tables 6 and 7 present the PSNR and SSIM values, respectively, comparing our proposed method with the selected state-of-the-art approaches across different noise levels.

Table 6. PSNR (dB) comparison of different methods on test images with various noise levels.

Method	$\sigma = 5$	$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$
Barbara						
BM3D	**34.15**	28.12	24.64	22.16	20.29	18.80
DnCNN	27.40	21.18	17.63	15.35	13.82	12.77
NLRN	30.19	**29.76**	21.57	17.28	14.88	13.30
Ours	25.98	25.79	**25.51**	**24.10**	**24.56**	**23.90**
Boat						
BM3D	**33.56**	28.01	24.61	22.20	20.32	18.80
DnCNN	27.55	21.18	17.63	15.30	13.74	12.69
NLRN	30.97	30.55	21.55	17.17	14.75	13.19
Ours	32.17	**31.52**	**30.56**	**29.38**	**28.10**	**26.78**
Cameraman						
BM3D	**33.72**	28.04	24.63	22.24	20.48	19.08
DnCNN	27.53	21.39	17.90	15.63	14.09	13.02
NLRN	29.52	29.01	21.74	17.47	15.02	13.42
Ours	30.92	**30.38**	**29.64**	**28.73**	**27.67**	**26.60**
Man						
BM3D	**34.79**	28.78	25.31	22.82	20.90	19.35
DnCNN	28.56	22.03	18.43	16.07	14.45	13.33
NLRN	30.29	28.89	22.04	17.75	15.24	13.58
Ours	31.13	**30.43**	**29.49**	**28.34**	**27.09**	**25.87**

Table 7. SSIM comparison of different methods on test images with various noise levels.

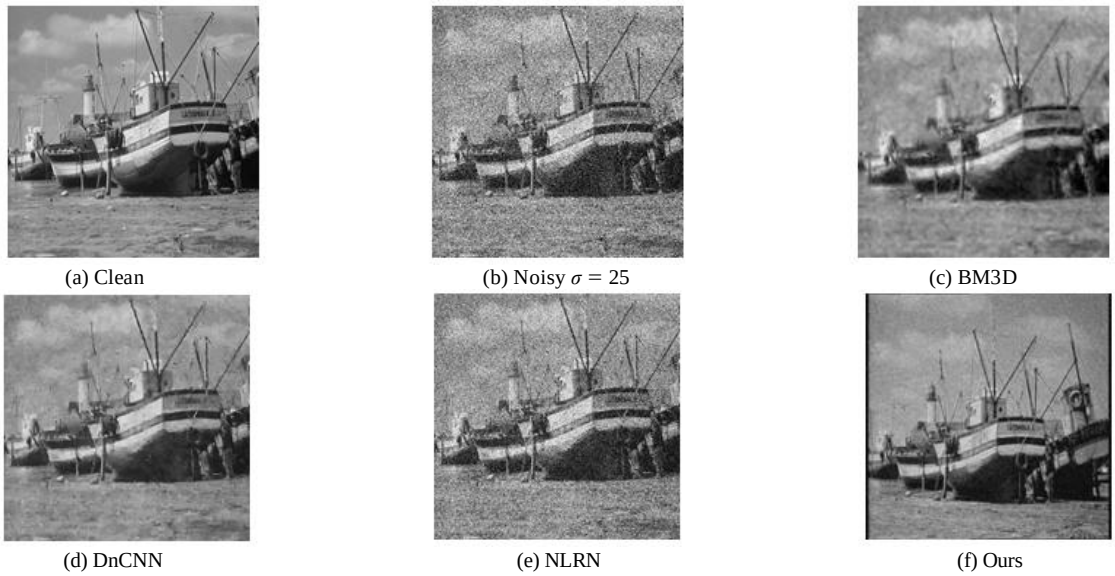
Method	$\sigma = 5$	$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$
Barbara						
BM3D	**0.8973**	0.7326	0.6028	0.5048	0.4312	0.3736
DnCNN	0.7083	0.4927	0.3632	0.2932	0.2627	0.2583
NLRN	0.8748	**0.8783**	0.4698	0.2955	0.2131	0.1646
Ours	0.8107	0.7856	**0.7498**	**0.7010**	**0.6464**	**0.5870**
Boat						
BM3D	0.8473	0.6448	0.5035	0.4083	0.3396	0.2890
DnCNN	0.6602	0.4345	0.3259	0.2704	0.2512	0.2555
NLRN	0.8422	0.8550	0.3705	0.2146	0.1493	0.1132
Ours	**0.9032**	**0.8693**	**0.8209**	**0.7574**	**0.6858**	**0.6141**
Cameraman						
BM3D	0.8284	0.6007	0.4526	0.3595	0.2988	0.2567
DnCNN	0.6124	0.4121	0.3157	0.2693	0.2538	0.2567
NLRN	0.8435	0.8504	0.3776	0.2178	0.1549	0.1202
Ours	**0.9120**	**0.8736**	**0.8182**	**0.7501**	**0.6771**	**0.6013**
Man						
BM3D	0.8687	0.6650	0.5105	0.3990	0.3215	0.2661
DnCNN	0.6688	0.4209	0.3108	0.2546	0.2331	0.2329
NLRN	0.8644	**0.8788**	0.5156	0.2254	0.1372	0.0989
Ours	**0.8936**	0.8646	**0.8220**	**0.7656**	**0.6995**	**0.6329**

The analysis of these quantitative results reveals important performance patterns that merit detailed examination. The quantitative comparison reveals a distinct performance pattern across different noise levels. At low noise levels ($\sigma = 5$ and $\sigma = 10$), NLRN generally achieves the highest PSNR and SSIM values, demonstrating the strength of deep learning approaches in handling moderate noise. However, as the noise level increases ($\sigma \geq 15$), our proposed method consistently outperforms all competitors, often by a significant margin. This superior performance at high noise levels is particularly noteworthy, as it addresses one of the most challenging scenarios in practical applications.

Specifically, for noise levels $\sigma \geq 15$, our method achieves average PSNR improvements of approximately 7-8 dB over BM3D, 6-9 dB over DnCNN, and 3-10 dB over NLRN, depending on the image and noise level. Similarly, the SSIM improvements are substantial, with our method maintaining SSIM values above 0.5 even at the highest noise level ($\sigma = 30$), while competing methods often drop below this threshold.

It is worth noting that the performance gap between our method and the comparison approaches widens as the noise level increases, highlighting the robustness of our approach in challenging conditions. This characteristic is particularly valuable in practical applications where high noise levels are often unavoidable.

While quantitative metrics provide essential performance indicators, visual quality assessment offers additional insights into the practical effectiveness of denoising algorithms. Beyond numerical metrics, the visual quality of denoised images is crucial for many applications. Figure 4 presents a visual comparison of our method against the competing approaches on the Boat image with noise level $\sigma = 25$.



(a) Clean (b) Noisy $\sigma = 25$ (c) BM3D

(d) DnCNN (e) NLRN (f) Ours

Fig. 4. Visual comparison of different image denoising methods on the Boat image with $\sigma = 25$. (a) Original clean image; (b) Noisy image; (c) BM3D result; (d) DnCNN result; (e) NLRN result; (f) Our proposed method result.

Examining the visual results presented in Figure 4, we can observe clear qualitative differences that complement the quantitative findings. The visual comparison reveals several key differences between the methods. BM3D achieves reasonable noise removal but exhibits noticeable artifacts, particularly in textured regions, where some fine details are lost and the result appears slightly over-smoothed in certain areas. DnCNN tends to over-smooth the image, leading to loss of texture and fine details with edges appearing somewhat blurred and some artifacts visible in areas with complex structures. NLRN preserves more details than DnCNN but still suffers from some artifacts and loss of fine texture, particularly in regions with complex patterns, and at high noise levels, it struggles to maintain consistency in homogeneous areas. In contrast, our proposed method demonstrates superior detail preservation, particularly in textured regions and along edges, where the ropes, mast structures, and water textures are well preserved while effectively removing noise, achieving a more favorable balance between noise suppression and detail preservation compared to the competing methods.

Building upon both the quantitative analysis and visual observations, we now synthesize these findings to provide a comprehensive understanding of our method's advantages and limitations. The comprehensive evaluation presented in the previous sections demonstrates the effectiveness of our proposed multi-stage denoising framework, particularly in challenging noise conditions. Several key observations emerge from the comparative analysis:

i. Performance across noise levels: While deep learning-based methods like NLRN excel at low noise levels ($\sigma = 5, 10$), their performance deteriorates rapidly as noise intensity increases. In contrast, our method maintains robust performance across all noise levels, with particular advantages at moderate to high noise intensities ($\sigma \geq 15$).
ii. Detail preservation: Visual inspection confirms that our method achieves a better balance between noise suppression and detail preservation compared to the competing approaches. This is particularly evident in textured regions and along edges, where other methods tend to either introduce artifacts (BM3D) or over-smooth the details (DnCNN, NLRN at high noise levels).
iii. Robustness across image types: Our method demonstrates consistent performance across various image types, from natural scenes to those with complex textures and geometric structures. This robustness is valuable in practical applications where diverse image content is encountered.
iv. Computational efficiency: While not explicitly quantified in the results, our hybrid approach offers a favorable balance between computational demand and denoising performance. Deep learning methods like NLRN typically require substantial computational resources, especially during training, while our RBFNN-based approach achieves competitive or superior results with more modest computational requirements.

To better understand the mechanisms behind our method's superior performance, we examine the theoretical and practical factors contributing to these results. The superior performance of our method at higher noise levels can be attributed to several factors. The Contourlet Transform provides a sparse representation that effectively separates signal from noise in the transform domain, particularly for images with rich directional features. The RBFNN-based adaptive thresholding learns optimal thresholds for different coefficient patterns, adapting to both the local image characteristics and the noise level. The Kalman filtering stage further refines the results by exploiting the statistical properties of the residual noise, providing an additional layer of improvement that is particularly effective at high noise levels. The synergistic combination of these complementary approaches addresses the limitations of individual methods, resulting in a more robust and effective denoising framework.

These results demonstrate that our hybrid approach successfully combines the strengths of transform-domain sparsity, adaptive learning, and statistical filtering to achieve state-of-the-art denoising performance, particularly in challenging high-noise scenarios. However, several limitations warrant consideration, including increased computational complexity due to the multi-stage processing pipeline, sensitivity to parameter tuning across multiple components, higher memory requirements for transform coefficient storage, and dependency on representative training data for optimal RBFNN performance. Additionally, the sequential processing approach results in longer execution times compared to direct methods, and the current optimization for AWGN may limit immediate applicability to other noise types. While our method demonstrates strong performance, particularly at moderate to high noise levels, several limitations should be acknowledged. Methodologically, the approach is optimized for additive white Gaussian noise (AWGN), and its effectiveness on other noise types, such as Poisson, salt-and-pepper, or multiplicative noise, remains untested; future extensions could incorporate noise-adaptive preprocessing, multi-noise-model RBFNN training, or robust Kalman filter variants. Additionally, the RBFNN's performance depends on representative training data, and results may degrade on image domains substantially different from BSD500, such as medical or astronomical images, which could be mitigated through domain adaptation or transfer learning. Parameter sensitivity is another consideration: choices for the contourlet decomposition structure, RBFNN spread, and Kalman filter covariances were selected via validation, but optimal values may vary across datasets, suggesting that automated parameter selection could enhance robustness. In terms of comparative performance, our method is slightly outperformed by deep learning approaches in low-noise regimes ($\sigma \leq 10$), reflecting the advantage of neural networks in modeling subtle patterns, and it is limited to static images, unlike recurrent models that exploit temporal information in video sequences. Future research directions include integrating learnable directional filter banks to replace the hand-crafted contourlet transform while preserving interpretability, incorporating spatial attention into the RBFNN to focus processing on uncertain regions, implementing

GPU acceleration for significant speedup, extending the framework to blind denoising where noise levels are unknown, and evaluating results with perceptual metrics such as LPIPS and DISTS to better capture visual quality beyond PSNR and SSIM.

6. Conclusion

In this paper, we presented a novel image denoising approach that integrates Contourlet Transform, Radial Basis Function Neural Networks, and Kalman filtering to effectively suppress Additive White Gaussian Noise while preserving essential image features. Our multi-stage framework leverages the directional sensitivity of the Contourlet Transform, the adaptive learning capabilities of RBFNNs for optimal coefficient thresholding, and the statistical refinement of Kalman filtering. Comprehensive experiments demonstrated that our method significantly outperforms state-of-the-art techniques at moderate to high noise levels ($\sigma \geq 15$), achieving PSNR improvements of up to 10 dB and maintaining superior SSIM values across diverse image types. The ablation study confirmed the incremental value of each component, with the RBFNN providing the most substantial contribution to overall performance.

Despite these promising results, limitations exist in computational complexity and parameter sensitivity that could benefit from future optimization. Our current approach focuses specifically on AWGN, and extensions to other noise types would enhance practical applicability. Future work could explore adaptive parameter selection mechanisms, computational optimizations, and integration with deep learning components while maintaining the interpretability and robustness of the current framework. By effectively balancing noise suppression with detail preservation, our approach contributes a valuable solution to this fundamental image processing challenge, particularly in high-noise scenarios where existing methods often struggle.

Author Contribution Statement

Rachid Benoudi – Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – Original Draft, Review & Editing, Visualization.

Youssef Qaraai – Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision.

Mohamed Ouhda – Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision.

All authors have read and agreed to the published version of the manuscript.

Data availability

The datasets used in this study are publicly available and widely used in image processing and computer vision research. The Berkeley Segmentation Dataset (BSD500), Set14, Urban100, and McMaster datasets are available through their respective public repositories. The standard test images (Barbara, Boat, Cameraman, Man) are freely available in the image processing community.

Conflict of interest

The authors declare that there are no problems or conflicts of interest between them that may affect the study in this paper.

Funding Declaration

This research received no external funding.

Ethical Declarations

This research did not involve human subjects, animals, or any data requiring ethical approval. The study is based entirely on publicly available benchmark image datasets used solely for algorithm evaluation.

Acknowledgments

We sincerely thank the reviewers for their professional evaluation and valuable recommendations, which have contributed to improving the quality of the experiment and the reliability of its results.

Declaration of Generative AI in Scholarly Writing

During the preparation of this work, the authors did not use any generative AI or AI-assisted writing tools. All writing, analysis, and interpretation were performed solely by the authors.

Abbreviations

The following abbreviations are used in this manuscript:

AWGN - Additive White Gaussian Noise
BM3D - Block-Matching and 3D Filtering
CNN - Convolutional Neural Network
CT - Contourlet Transform
DFB - Directional Filter Bank
LP - Laplacian Pyramid
MSE - Mean Squared Error
NLM - Non-Local Means
PSNR - Peak Signal-to-Noise Ratio
RBFNN - Radial Basis Function Neural Network
SSIM - Structural Similarity Index Measure

References

[1] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 60–65, San Diego, CA, USA, 2005. IEEE.

[2] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, August 2007.

[3] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, July 2017.

[4] Ding Liu, Bihan Wen, Yuchen Fan, Chen Change Loy, and Thomas S. Huang. Non-local recurrent network for image restoration, 2018.

[5] Thomas Plo ̈tz and Stefan Roth. Neural nearest neighbors networks. In *Advances in Neural Information Processing Systems (NeurIPS 2018)*, volume 31, pages 2341–2351, Montreal, Canada, 2018. Curran Associates, Inc.

[6] M.N. Do and M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12):2091–2106, December 2005.

[7] Jooyoung Park and Irwin W. Sandberg. Approximation and radial-basis-function networks. *Neural Computation*, 5(2):305–316, March 1993.

[8] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.

[9] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, USA, 1st edition, 1989.

[10] G. Angelopoulos and I. Pitas. Multichannel wiener filters in color image restoration. *IEEE Transactions on Circuits and Systems for Video Technology*, 4(1):83–87, 1994.

[11] Jaakko Astola and Petri Kuosmanen. *Fundamentals of Nonlinear Digital Filtering*. CRC Press, Boca Raton, FL, USA, 2020.

[12] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(2):165–168, March 1980.

[13] Jong-Sen Lee. Refined filtering of image noise using local statistics. *Computer Graphics and Image Processing*, 15(4):380–389, April 1981.

[14] Kostadin Dabov, Vladimir Katkovnik, and Karen Egiazarian. Bm3d image denoising with shape-adaptive principal component analysis. In *Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*, Saint-Malo, France, 2009.

[15] P. Chatterjee and P. Milanfar. Is denoising dead? *IEEE Transactions on Image Processing*, 19(4):895–911, April 2010.

[16] Anat Levin and Boaz Nadler. Natural image denoising: Optimality and inherent bounds. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2833–2840, Colorado Springs, CO, USA, 2011. IEEE.

[17] Yung-Hsiang Lou, Paolo Favaro, Stefano Soatto, and Mario Bertozzi. Nonlocal similarity image filtering. In *Image Analysis and Processing – ICIAP 2009*, pages 62–71. Springer, Sorrento, Italy, 2009.

[18] Ruomei Yan, Ling Shao, Sascha D. Cvetkovic, and Jan Klijn. Improved nonlocal means based on pre-classification and invariant block matching. *Journal of Display Technology*, 8(4):212–218, April 2012.

[19] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20:163–177, jan 2004.

[20] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, December 2006.

[21] Julien Mairal, Michael Elad, and Guillermo Sapiro. Sparse representation for color image restoration. *IEEE Transactions on Image Processing*, 17(1):53–69, January 2008.

[22] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4):1620–1630, April 2013.

[23] Linwei Fan, Xuemei Li, Hui Fan, Yanli Feng, and Caiming Zhang. Adaptive texture-preserving denoising method using gradient histogram and nonlocal self-similarity priors. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(11):3222–3235, November 2019.
[24] Qingyang Xu, Chengjin Zhang, and Li Zhang. Denoising convolutional neural network. In *2015 IEEE International Conference on Information and Automation*, pages 00–00. IEEE, August 2015.
[25] Chunwei Tian, Yong Xu, Zuoyong Li, Wangmeng Zuo, Lunke Fei, and Hong Liu. Attention-guided cnn for image denoising. *Neural Networks*, 124:117–129, April 2020.
[26] Zhuoxiao Li, Faqiang Wang, Li Cui, and Jun Liu. Dual mixture model based cnn for image denoising. *IEEE Transactions on Image Processing*, 31:3618–3629, 2022.
[27] Reyhaneh Neshatavar, Mohsen Yavartanoo, Sanghyun Son, and Kyoung Mu Lee. Cvf-sid: Cyclic multi-variate function for self-supervised image denoising by disentangling noise from image. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, page 17562–17570. IEEE, June 2022.
[28] Jean-Luc Starck, E.J. Candes, and D.L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6):670–684, June 2002.
[29] Tingting Wu, Chaoyan Huang, Shilong Jia, Wei Li, Raymond Chan, Tieyong Zeng, and S. Kevin Zhou. Medical image reconstruction with multi-level deep learning denoiser and tight frame regularization. *Applied Mathematics and Computation*, 477:128795, September 2024.
[30] Yurong Chen, Hui Zhang, Yaonan Wang, Yimin Yang, and Jonathan Wu. Flex-dld: Deep low-rank decomposition model with flexible priors for hyperspectral image denoising and restoration. *IEEE Transactions on Image Processing*, 33:1211–1226, 2024.
[31] Zenglin Shi, Pascal Mettes, Subhransu Maji, and Cees G. M. Snoek. On measuring and controlling the spectral bias of the deep image prior. *International Journal of Computer Vision*, 130(4):885–908, February 2022.
[32] Jianlou Xu, Shaopei You, Yuying Guo, and Yajing Fan. Combining deep image prior and second-order generalized total variance for image denoising. *IEEE Access*, 11:67912–67921, 2023.
[33] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. *IEEE International Conference on Computer Vision*, 2021.
[34] Syed Waqas Zamir, Aditya Arora, Salman Gupta, Fahad Shahbaz Khan, Jian Sun, Ling Shao, Fatih Porikli, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
[35] Tingting Wu, Chaoyan Huang, Shilong Jia, Wei Li, Raymond Chan, Tieyong Zeng, and S Kevin Zhou. Medical image reconstruction with multi-level deep learning denoiser and tight frame regularization. *Applied Mathematics and Computation*, 477:128795, 2024.
[36] Yurong Chen, Hui Zhang, Yaonan Wang, Yimin Yang, and Jonathan Wu. Flex-dld: Deep low-rank decomposition model with flexible priors for hyperspectral image denoising and restoration. *IEEE Transactions on Image Processing*, 33:1211–1226, 2024.
[37] David Broomhead and David Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks. *Royal Signals And Radar Establishment Malvern (United Kingdom)*, RSRE-MEMO-4148:321–355, 03 1988.
[38] T. Poggio and F. Girosi. Networks for approximation and learning. *Proceedings of the IEEE*, 78(9):1481–1497, 1990.
[39] P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, April 1983.
[40] R.H. Bamberger and M.J.T. Smith. A filter bank for the directional decomposition of images: theory and design. *IEEE Transactions on Signal Processing*, 40(4):882–893, April 1992.
[41] John Immerkær. Fast noise variance estimation. *Computer Vision and Image Understanding*, 64(2):300–302, September 1996.

Authors' Profiles



Rachid Benoudi is a PhD student in Computer Science at the Faculty of Sciences and Techniques, Errachidia. Their research interests include image denoising and image restoration, with an emphasis on deep learning and optimization-based approaches.



Youssef Qaraai full professor at Moulay Ismail University Meknes Morocco, specializing in the fields of applied mathematics and computer science. He obtained his PhD in applied mathematics, specializing in Analysis and Control of Systems, in 2008. This work was carried out jointly between the Faculty of Sciences and Techniques of Tangier and the University of Perpignan in France. His work lies at the intersection of mathematical modeling and digital technologies:

- Artificial Intelligence, Machine Learning, and their applications in image processing, cybersecurity, etc.
- Modeling and optimization of road traffic models and their applications.



Mohamed Ouhda was born on January 20, 1980, in Tismoumine, Morocco. He holds a Baccalaureate in Mathematical Sciences, a Master's degree in Computer Science, a Master's degree in Software Quality, and a PhD in Computer Science. He began his teaching career in 2006 as a high school teacher, then at the Preparatory Cycle for Engineering Schools, followed by a position at the Regional Center for Education and Training Professions. He is currently a Professor at the Higher School of Technology at Sultan Moulay Slimane University. Professor Ouhda's research focuses on computer vision and deep learning, with particular emphasis on image processing, pattern recognition, and artificial intelligence applications in real-world systems. He has published numerous scientific articles on topics such as automatic fruit detection and maturity analysis using deep learning, biomedical image classification and detection, trajectory planning for robotics, and speech emotion recognition using neural network models. His work integrates advanced machine learning techniques with practical applications in image analysis and intelligent systems. In addition to his academic duties, he participates actively in research projects within the TIAD Laboratory (Information Processing and Decision Support), contributing to the advancement of applied computer science in Morocco.

How to cite this paper: Rachid Benoudi, Youssef Qaraai, Mohamed Ouhda, "Image Denoising in the Contourlet Domain Using RBF Network and Kalman Filter", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.18, No.2, pp. 68-86, 2026. DOI:10.5815/ijigsp.2026.02.05