fractal and fractional

MDPI

Article

DGGNets: Deep Gradient-Guidance Networks for Speckle Noise Reduction

Li Wang [1], Jinkai Li [2], Yi-Fei Pu [1,*], Hao Yin [1] and Paul Liu [3]

[1] College of Computer Science, Sichuan University, Chengdu 610064, China; wangli5@stu.scu.edu.cn (L.W.); yinhao@scu.edu.cn (H.Y.)
[2] Bio-Computing Research Center, Harbin Institute of Technology, Shenzhen 518055, China; 20b951007@stu.hit.edu.cn
[3] Stork Healthcare, Chengdu 610064, China; paul@storkhealthcare.com
* Correspondence: puyifei@scu.edu.cn

Abstract: Speckle noise is a granular interference that degrades image quality in coherent imaging systems, including underwater sonar, Synthetic Aperture Radar (SAR), and medical ultrasound. This study aims to enhance speckle noise reduction through advanced deep learning techniques. We introduce the Deep Gradient-Guidance Network (DGGNet), which features an architecture comprising one encoder and two decoders—one dedicated to image recovery and the other to gradient preservation. Our approach integrates a gradient map and fractional-order total variation into the loss function to guide training. The gradient map provides structural guidance for edge preservation and directs the denoising branch to focus on sharp regions, thereby preventing over-smoothing. The fractional-order total variation mitigates detail ambiguity and excessive smoothing, ensuring rich textures and detailed information are retained. Extensive experiments yield an average Peak Signal-to-Noise Ratio (PSNR) of 31.52 dB and a Structural Similarity Index (SSIM) of 0.863 across various benchmark datasets, including McMaster, Kodak24, BSD68, Set12, and Urban100. DGGNet outperforms existing methods, such as RIDNet, which achieved a PSNR of 31.42 dB and an SSIM of 0.853, thereby establishing new benchmarks in speckle noise reduction.

Keywords: speckle noise reduction; fractional-order total variation; gradient preservation; deep learning

Citation: Wang, L.; Li, J.; Pu, Y.-F.; Yin, H.; Liu, P. DGGNets: Deep Gradient-Guidance Networks for Speckle Noise Reduction. *Fractal Fract.* **2024**, *8*, 666. https://doi.org/10.3390/fractalfract8110666

Academic Editor: Norbert Herencsar

Received: 28 August 2024
Revised: 7 November 2024
Accepted: 13 November 2024
Published: 15 November 2024

1. Introduction

Coherent imaging systems, such as underwater sonar [1], Synthetic Aperture Radar (SAR) [2], and medical ultrasound systems [3], have been widely studied in artificial intelligence, as shown in Figure 1. This figure illustrates the specific application scenarios and objectives of the algorithm. However, their image quality is highly influenced by the speckle noise, which is locally correlated. The speckle noise can hide the finer features of the image and reduce accuracy in industrial practical applications. Therefore, speckle noise reduction is very important in enhancing and segmenting coherent images. It involves sharpening and connecting edges while smoothing speckle regions without causing image blurring. Furthermore, to make the recovered image structure clear, the methods must increase the overall contrast but not the graininess of speckles. In other words, speckle noise reduction aims to augment structural contrast while reducing speckle contrast. The motivation of this study is to address the challenges of preserving structural details and avoiding over-smoothness in coherent image denoising. Traditional methods struggle with balancing noise reduction and edge preservation, leading to either inadequate noise suppression or the loss of important structural features. This study aims to overcome these limitations by introducing a novel approach that combines gradient information with deep learning techniques to achieve superior edge preservation and reduced over-smoothing. Thus, speckle reduction methods must fulfill essential functions: speckle noise reduction, edge preservation, boundary enhancement, and texture recovery.

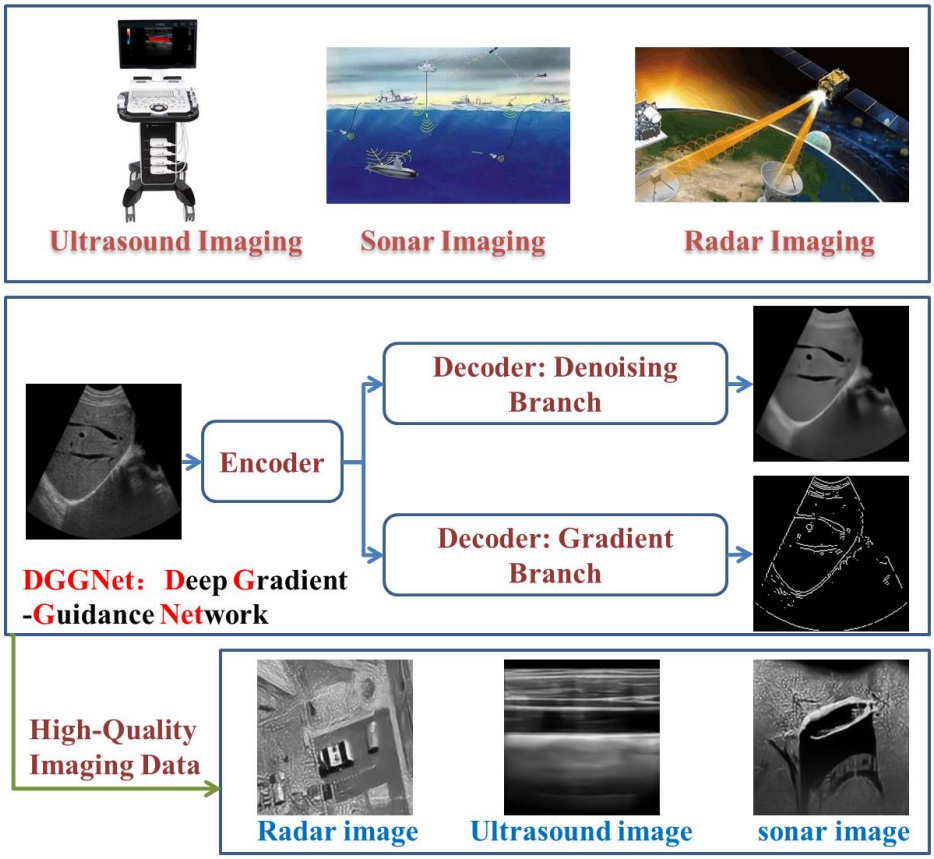

Figure 1. System architecture of a speckle noise reduction system.

Various denoising methods aim to address noise in coherent images [4]. Traditional approaches, including spatial domain filters, diffusion filters, Non-Local Mean (NLM) filters, transform domain techniques, and hybrid filters, have made significant strides in image denoising [5]. Spatial domain filter methods [6,7] assume multiplicative speckle noise, utilizing moving filter windows based on local statistical properties to remove noise. However, they may blur the recovered image due to lacking global structure information. Diffusion filter techniques, e.g., the Speckle-Reducing Anisotropic Diffusion filter (SRAD) [8], aim to enhance image quality but tend to introduce stair-step effects, resulting in information loss and over-filtering. Non-Local Mean filters combine features of the same cluster. However, these methods are complex and sensitive to patch quality, and they can potentially cause blurring when patches are dissimilar. Transform domain techniques denoise images using transformed domain properties, e.g., wavelet-based thresholding [9,10], and tend to introduce artifacts in the denoising results. Zhou et al. [11] proposed a degradation model based on the nonlinear transformation to adjust the intensity of SAR image pixel values. Bi et al. [12] proposed a mixed-order image denoising algorithm containing fractional-order and high-order regularization terms, which effectively suppresses the staircase effect generated by the TV model and its variants while better preserving the edges and details of the image. Hybrid filter approaches combine methods but suffer from high complexity, limiting the effectiveness of traditional denoising approaches.

In recent years, deep learning techniques have shown significant progress in image denoising. The multi-scale-based method [13] and CNN-based method [14] have demonstrated their advantages. Liang et al. [15] pioneered deep learning for image denoising. Denoising Convolutional Neural Network (DnCNN) [16] is a universal method but does not consider underlying structures and textures. Residual Image Denoising Network (RIDNet) [17] recovers the image by a single-stage method with the global residual. Attention mechanisms, e.g., Deep Residual Attention Network (DRAN) [18] and Mixed-Attention

Residual U-Net (MARU) [19], show their promise in medical image denoising. The Multi-Scale Attention Network (MSANN) [20] uses a multiscale attention module to denoise. However, their effectiveness on real ultrasound images is uncertain, and denoised images often exhibit overly smooth structures. Performance on real coherent image datasets is yet to be confirmed.

In this paper, we propose a two-branch denoising method to realize edge preservation and over-smoothness alleviation. Since the gradient map reveals the region where sharpness is in the image, it is a powerful implementation to guide image restoration; we design a dual decoder architecture in the proposed network that combines an image denoising branch and a gradient branch. The gradient branch works in two ways: (i) the gradient branch provides more structural information to the denoising branch, and (ii) the multi-level gradient feature can make the denoising branch pay more attention to the sharp region, which can contribute to the structure-preserving and relieve over-smoothness. Thus, our proposed DGGNet consists of one encoder and two novel decoders (one is for image recovery, and the other is for edge preservation). In addition, inspired by the fractional total variation that can find the fine boundary in signal processing and image processing [21], we introduced it in our network to preserve texture details. Our method is designed to improve both edge preservation and over-smoothing reduction. In the subsequent sections, we present experimental results and ablation studies that validate the effectiveness of the proposed gradient branch and fractional total variation.

The main contributions of this work are summarized in the following:

- We propose a novel structure-preserving denoising method that utilizes a gradient branch to guide the recovery of coherent images. This method addresses the challenge of restoring images from heavy speckle noise while avoiding over-smoothing. To our knowledge, we are the first to leverage gradient maps to guide speckle noise reduction.
- We introduce a fractional total variational loss function specifically designed to prevent detail ambiguity and over-smoothing, which are common issues with integral order methods. This approach effectively preserves rich texture and detail information in the denoised images. The effectiveness of this loss function is demonstrated through comprehensive ablation experiments.
- We conduct extensive experiments on established real coherent image datasets. Our method is compared with existing approaches, and the results show that it achieves state-of-the-art performance in both quantitative metrics and visual quality.

This paper is organized as follows: Section 2 describes our proposed method. Section 3 presents the experimental results on synthetic and realistic images. The conclusion is presented in Section 4.

2. Related Work

2.1. Coherent Imaging Denoising

Various efforts have been proposed to remove the noise in coherent images [4]. Traditional methods, which can be classified into spatial domain filters, diffusion filters, Non-Local Mean (NLM) filters, transform domain techniques, and hybrid filter approaches, have achieved significant progress in image denoising [5]. Spatial domain filter methods [6,7] assume that speckle noise is multiplicative with respect to the noise-free image. In spatial domain filter methods, the pixel in the noise-free image is estimated by a moving filter window whose weighting coefficients are designed based on statistical properties. However, these methods only use local information, which can lead to the blurring of the recovered image due to the loss of global structure. Furthermore, spatial domain methods may struggle to preserve fine details and textures, potentially leading to suboptimal results in complex scenarios. Diffusion filter techniques are widely used to improve image quality, for example, the Speckle-Reducing Anisotropic Diffusion filter (SRAD) [8]. However, it is important to note that diffusion filters can result in stair-step effects and a loss of information while diffusing, which may cause over-filtering. Moreover, diffusion filters often

face challenges in maintaining edge details, which can impact the overall image quality. In addition, most of these filters cannot be used for real-time processing.

The Non-Local Mean filters include Non-Local Mean (NLM) [22], the Optimal Bayesian NLM (OBNLM) filter [23], and the Non-Local Low-Rank Framework (NLLRF) [7]. These techniques typically combine various features from the same cluster and exhibit high complexity. They are highly dependent on the quality of the selected patches; if the chosen patches are dissimilar, the output may end up being blurred. Additionally, the computational complexity of these methods can be prohibitive, especially for large images or real-time applications. Low-rank-based methods attempt to approximate the underlying low-rank matrix from the noisy input [24]. However, these methods perform well only when the degraded image adheres to the low-rank characteristic. Transform domain techniques denoise images by leveraging properties of the transformed domain image, such as wavelet-based thresholding techniques [9,10]. Despite their effectiveness, these techniques can introduce artifacts, which can degrade the visual quality of the denoised images. Hybrid filter approaches combine two or more methods for image denoising but suffer from high complexity [4]. The combination of multiple methods often results in increased computational requirements and can complicate the implementation. All of the above limitations have restricted the effectiveness of traditional image-denoising approaches.

In recent years, deep learning techniques have been widely used for image denoising, resulting in significant advancements. Some effective speckle-reduction methods based on deep learning have been developed, including super-resolution via image-adapted denoising CNNs [25], channel and space attention neural networks [26], and grouped multi-scale networks [13]. These techniques have shown various advantages in their respective studies.

Liang et al. [15] were the first to apply deep learning networks to image denoising. Subsequently, Zhang et al. [16] proposed a feed-forward Denoising Convolutional Neural Network (DnCNN), which is a universal method for image noise reduction and super-resolution. DnCNN restores the clean image by estimating the noise present in the input image. However, it does not account for underlying image structures and textures, as highlighted by Anwar et al. [17]. To address the limitations of DnCNN, Anwar et al. proposed a blind Real Image Denoising Network (RIDNet), which is a one-stage method that achieves good results in natural image denoising. Following this, attention mechanisms have been introduced in deep learning denoising methods, leading to very promising results. For instance, Sharif et al. [18] combined attention mechanisms with spatially refined residual features to address multidisciplinary medical image denoising in the Deep Residual Attention Network (DRAN). Similarly, Lan et al. [19] proposed a Mixed-Attention Residual U-Net (MARU), which uses a mixed-attention mechanism for ultrasound image denoising. A Multi-Scale Attention-guided Neural Network (MSANN) [20] uses a multiscale attention module to obtain the spatial distribution of speckle noise for processing complex noise.

However, while these methods have been shown to work well with synthetic data, the effectiveness of denoising real ultrasound images remains uncertain. All denoised images generated by these deep-learning techniques often suffer from overly smooth structures. Furthermore, their performance on real ultrasound image datasets has yet to be thoroughly validated. Unlike conventional approaches, which tend to cause over-smoothing, our proposed DGGNet is designed to strike a balance between reducing speckle noise and avoiding excessive smoothing.

2.2. Edge Preservation in Denoising

Image edges are the boundaries between different regions or objects in an image, and they contain significant information about the image structure and semantics. However, image edges are often corrupted by noise, blur, or compression artifacts, which degrade image quality and affect subsequent analysis or processing. Image edge preservation is a crucial technique in image processing, especially for applications requiring high-quality

images with clear and undistorted edges, such as medical imaging, remote sensing, and computer vision.

There are several methods for denoising images while preserving their edges. One such method is the Non-Local Means (NLM) filter [22]. This method replaces each pixel with the weighted average of all pixels in the image, with the weights based on the similarity between the patches centered at the pixels. Another method is the guided filter [27]. This method applies a guided image filter to the noisy image, with the guidance provided by a smoothed version of the noisy image, which helps preserve edges while reducing noise. Finally, the anisotropic diffusion filter [8] applies a diffusion process to the image, with the diffusion tensor based on the gradient of the image. These methods aim to reduce noise while preserving edges, corners, and other sharp structures in images.

There are also several edge-preserving image-denoising methods based on deep learning. One such method is MSANN [20]. It can effectively remove speckle noise using an intelligent computing-enabled multi-scale attention-guided neural network. Another method is Adaptive Thresholding-based DWT [28], which utilizes a modified deep-structured architecture. This approach applies adaptive thresholding in the Discrete Wavelet Transform (DWT) domain and a modified Deep Convolutional Neural Network (DCNN) to protect image edges. Unlike other approaches influenced by the quality and diversity of the training data, our training process extends beyond synthetic image datasets to incorporate real ultrasound image datasets. This ensures a more comprehensive and robust model that can better generalize to real-world scenarios.

2.3. Fractional-Order Total Variation

Fractional calculus has been successfully applied to signal processing and image processing because of its weak singularity, long-term memory, and non-locality [29]. Pu introduced fractional calculus to image processing, leading to the development of a fractional differential mask [30]. This mask is capable of preserving low-frequency contour features in smooth areas of an image while enhancing high-frequency edges and textural details where there are significant variations in grey levels.

Total variation regularization is a well-known sparse representation technique for image denoising. However, it relies on the assumption of piecewise smooth signals, which can lead to over-smoothing. Total variation regularization struggles with preserving intricate details like textures due to this limitation. Various approaches incorporating higher-order regularization terms have been proposed to address this issue; however, they often introduce artifacts resembling speckle noise [31]. To strike a balance between over-smoothing and minimizing noise, Dali et al. [32] proposed integrating a fractional-order regularization term into image denoising. Due to its superior ability to capture subtle structures, fractional-order regularization excels at preserving valuable information in complex patterns. We incorporate this fractional regularization term into our loss function for speckle noise reduction to mitigate detail ambiguity and over-smoothing.

An efficient matrix approximation method for fractional-order derivatives was proposed by [33], which results in significant simplification in numerical solutions.

Consider the provided signal $f(x)$ obtained by sampling its continuous counterpart on a uniformly spaced grid with grid size Δh. Consequently, the discrete representation is given by $f(x_i) = f(i\Delta h)$, for $i = 0, 1, \ldots, m$. In image processing applications, the grid size Δh is commonly set to one, simplifying the description to $\Delta h = 1$. Utilizing the Grünwald–Letnikov (GL) fractional derivative definition [33], we can express the discrete formula for the fractional-order derivative of the digital signal as follows:

$$a^{G-L}D_x^v \approx \sum_{j=0}^{m} w_j^v f\left(x_{m-j}\right), \qquad \begin{array}{ll} m = 0, & j = 0 \\ m = 1, & j = 0, 1 \\ m = 2, & j = 0, 1, 2 \\ \ldots & \end{array} \tag{1}$$

where $w_j^v = (-1)^j \binom{v}{j}$ represent the coefficients. The coefficients can also be obtained recursively from

$$w_0^v = 1, \quad w_j^v = (1 - \frac{v+1}{j}) w_{j-1}^v, \quad j = 1, 2, \ldots. \tag{2}$$

Using the matrix approximation method, Equation (4) can be rewritten as follows:

$$D_x^v \overrightarrow{f(x)} \approx \psi \overrightarrow{f(x)} \tag{3}$$

where $\overrightarrow{f(x)} = [f(x_0), f(x_1), \ldots, f(x_m)]^T$ and ψ is a matrix defined as

$$\psi = \begin{bmatrix} w_0^v & 0 & \ldots & 0 \\ w_1^v & w_0^v & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ w_m^v & w_{m-1}^v & \ldots & w_0^v \end{bmatrix}. \tag{4}$$

When $v = 1$, the matrix ψ takes the form of a sparse banded matrix, consisting solely of two diagonals: the main diagonal and a single lower diagonal. Consequently, the gradient information is determined using only two data points. In cases where v is not an integer, the matrix ψ transforms into a lower triangular matrix. Notably, all data points preceding the kth point are utilized to compute the fractional-order derivative of the point. This enduring reliance on past data points signifies a crucial aspect of fractional differentiation, distinguishing it significantly from its integer-order counterpart. Due to its long-term memory, fractional differentiation is an essential tool for modeling characteristic phenomena in various applications.

3. Proposed Method

The proposed Deep Gradient-Guidance Networks (DGGNet) for coherent image denoising can be described as follows:

$$\hat{y}, \hat{e} = DGGNet(u), \tag{5}$$

where $DGGNet$ is the proposed network, u is the input of noisy coherent image, $\hat{y}$ is the output of the denoising result, and $\hat{e}$ is the output of the gradient of the image.

The overall architecture of the proposed network is illustrated in Figure 2. This figure shows that the output of the encoder is directed to two separate decoder branches: one for image denoising and the other for gradient extraction.

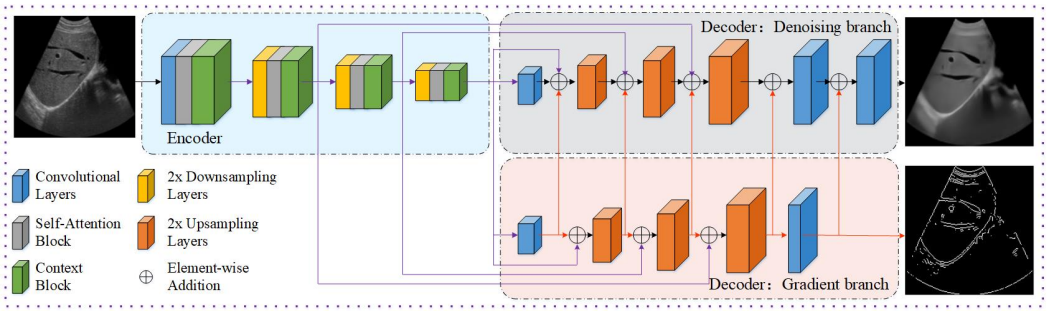

Figure 2. The network structure of the proposed DGGNet. The DGGNet consists of one encoder and two decoders (one decoder works for the denoising branch, and the other works for the gradient branch). The gradient branch guides the denoising branch by fusing gradient information to enhance structure preservation.

3.1. Denoising Branch

The denoising branch is designed to effectively reconstruct a noise-free image while preserving structural details. It utilizes a series of convolutional layers and context blocks, which incorporate multi-scale feature extraction. Each context block employs multiple dilated convolutions to capture information across various scales, allowing the network to better understand the structure and texture of the input image. The output of these context blocks is concatenated and fused, ensuring that important details are maintained while noise is reduced.

To mitigate the problem of over-smoothing, the denoising branch integrates gradient information from the gradient branch. The gradients highlight areas of sharpness and smoothness, enabling the denoising branch to focus on restoring fine details. By imposing gradient constraints, the denoising branch can effectively reconstruct sharp structures and reduce the blurring commonly encountered in traditional denoising approaches.

3.2. Gradient Branch

The gradient branch aims to estimate gradient maps from the noisy image. The gradient map, denoted as $Map(u)$, is derived from the differences between adjacent pixels, which can be expressed mathematically as

$$\begin{aligned} &u_x(x,y) = u(x+1,y) - u(x-1,y), \\ &u_y(x,y) = u(x,y+1) - u(x,y-1), \\ &\nabla u(x,y) = \big(u_x(x,y), u_y(x,y)\big), \\ &Map(u) = \parallel \nabla u \parallel_2, \end{aligned} \tag{6}$$

where $Map(.)$ extracts the gradient map, providing essential information about the image's edges. The gradient extraction can be efficiently implemented using convolutional layers with fixed kernels.

3.3. Integration of Branches

The integration of the gradient and denoising branches forms a powerful framework for image processing. The gradient branch acts as a guiding mechanism, enhancing the denoising branch's ability to focus on sharp regions of the image. This dual-branch structure allows for effective noise reduction without compromising on detail preservation.

The denoising branch also employs a residual architecture with skip connections, which helps retain fine image details by allowing the gradient and contextual features to be directly incorporated into the reconstruction process. Additionally, self-attention mechanisms are employed within both branches to capture long-range dependencies and contextual relationships, further enhancing the network's capability to process images coherently.

3.4. Dilation Rates in Context Blocks

To improve the network's performance and generalization capabilities, we incorporate five dilation rates in the context blocks of DGGNet. The dilation rates are set to 1, 2, 4, 8, and 16, with corresponding padding values configured to ensure that the dimensions of output features are consistent with those of input features. This design allows the model to capture a wider context without losing resolution. Finally, the features extracted from different dilation rates are concatenated along the channel dimension and mapped back to the original dimensionality using a convolutional layer.

3.5. Workflow

The algorithm workflow starts by dividing the dataset into training, validation, and testing sets. The testing set undergoes preliminary structural testing to ensure data integrity. The training set is used to train the DGGNet network, with performance evaluated on the validation set after each training iteration to monitor and adjust model parameters. After the model is fully trained, the testing set is used for final evaluation, producing

denoised images as the final output. Figure 3 provides a detailed flow diagram of the proposed algorithm.

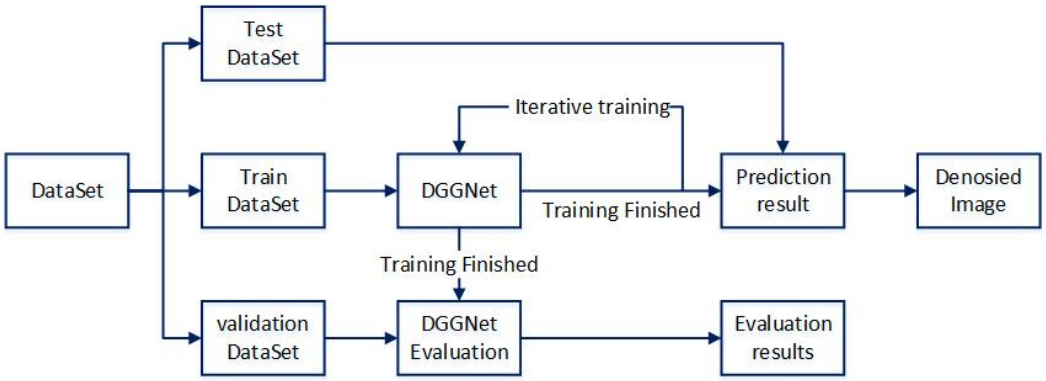

Figure 3. The flow diagram of the proposed DGGNet.

3.6. Fractional Total Variation Regularization

The Grünwald–Letnikov method of fractional calculus is primarily used in signal processing. Fractional-order differentials can more flexibly capture details and texture information in images. Compared to integer-order derivatives, fractional-order derivatives are better suited to handle images with non-integer statistical characteristics. Therefore, applying fractional-order differential to image denoising allows for better preservation of image details. We adopt the fractional-order differential expression based on Grünwald–Letnikov. From [30], we obtain the fractional differential as this formula:

$$\frac{d^v}{dx^v} f(x)_{G-L} \cong \frac{x^{-v} N^v}{\Gamma(-v)} \sum_{k=0}^{N-1} \frac{\Gamma(k-v)}{\Gamma(k+1)} f\left(x + \frac{vx}{2N} - \frac{kx}{N}\right), \tag{7}$$

where $f(x)$ is a differ-integrable function, the duration of $f(x)$ is $[a, x]$, and v is any real number (including a fraction) that indicates the order of the derivation. $\Gamma(v)$ is a gamma function. We use this formula as our fractional filter for preserving texture in coherent images.

Based on the filtering function, the fractional filter acts as an attenuation function in the low-frequency range ($0 < \omega < 1$). In the $\omega > 1$ range, it boosts amplitude, and as the fractional order v increases, the enhancement intensifies. Fractional-order differentiation can effectively preserve texture in high-frequency image denoising, ensuring enhanced amplitude in that spectrum. Thus, the fractional total variation regularization can be utilized for denoising images while preserving both details and edges within the image.

Let $f(x, y) = u(x, y) + \eta(x, y)$ be the observed noisy image, where (x, y) is the location with a rectangular image domain $\Omega \in R^2$, and η is noise. The $\mathrm{TV_p^v} - \mathrm{L}^2$ denoising model can estimate the desired clean image $u(x, y)$ by solving the following finite-dimensional optimization problem:

$$\hat{u} = \arg\min_u \left\{ E(u) = ||f - u||_2^2 + \lambda \mathrm{TV}_p^v(u) \right\}, \tag{8}$$

where $||f - u||_2^2$ is the data fidelity term, $TV_p^v(u)$ is the regularization term, and λ is the regularization parameter, which controls the level of smoothing.

In this work, we use the discrete definition of the regularization term $TV_p^v(u)$ to guide the training, which is given by the following formula:

$$\mathrm{TV}_p^v(u) = \sum_{n \in \Omega} ||\nabla^v u||_p, \tag{9}$$

where ∇^v is the fractional-order derivative operator, u is the noisy input image, and p is the order of the derivative operator.

Let u_x^v and u_y^v be linear operators corresponding to horizontal and vertical fractional-order derivatives. In practical work, we use the matrix approximate method [33] for the convenient implementation of fractional-order differentiation in the experiments. u_x^v and u_y^v in the matrix approximate method can be written in the following form:

$$\begin{cases} u_x^v = \nabla_x^v \cdot u \\ u_y^v = u \cdot \nabla_y^v, \end{cases} \tag{10}$$

where

$$\nabla_x^v = \begin{bmatrix} w_0^v & 0 & \dots & 0 \\ w_1^v & w_0^v & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ w_m^v & w_{m-1}^v & \dots & w_0^v \end{bmatrix}, \nabla_y^v = \begin{bmatrix} w_0^v & w_1^v & \dots & w_m^v \\ \vdots & w_0^v & \dots & w_{m-1}^v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_0^v \end{bmatrix}. \tag{11}$$

3.7. Loss Function

The proposed DGGNet uses three loss functions to supervise the training. $\mathcal{L}_{ID}$ works for image denoising, $\mathcal{L}_{GB}$ works for the gradient branch, and $\mathcal{L}_{FTV}$ works for artifact removal and overall image texture retention. The losses of $\mathcal{L}_{ID}$ and $\mathcal{L}_{GB}$ employ the $l1$ loss, while the $\mathcal{L}_{FTV}$ loss utilizes fractional total variation. They are defined as follows: assume that there are N training pairs $\{u_i, y_i\}_{i=1}^{i=N}$, where u is the noisy input image and y is the corresponding ground truth. $\hat{y}$ is the denoising results of the DGGNet, $\hat{e}$ is the gradient estimation results of the DGGNet, and e is the corresponding ground truth, which is applied by the Sobel filter [34]. The $\mathcal{L}_{ID}$ can be formulated as

$$\mathcal{L}_{ID} = \frac{1}{N} \sum_{i=1}^{N} ||\hat{y}_i - y_i||_1. \tag{12}$$

$\mathcal{L}_{GB}$ can be formulated as

$$\mathcal{L}_{GB} = \frac{1}{N} \sum_{i=1}^{N} ||\hat{e}_i - e_i||_1. \tag{13}$$

As discussed above, total variation regularization is an effective sparse representation technique for image denoising that can keep the details of the image texture while denoising. The fractional-order variational loss function $\mathcal{L}_{FTV}$ can be formulated as

$$\mathcal{L}_{FTV} = \sqrt{(\hat{y}_x^v)^2 + \left(\hat{y}_y^v\right)^2}. \tag{14}$$

Now, the total loss of our proposed DGGNet can be formulated in the following form:

$$\mathcal{L}_{total}(\Theta) = \mathcal{L}_{ID} + \lambda_{GB}\mathcal{L}_{GB} + \lambda_{FTV}\mathcal{L}_{FTV}, \tag{15}$$

where Θ denotes all of the learned parameters in DGGNet and λ_{GB} and λ_{FTV} are the balancing weights. Our aim with the proposed DGGNet is to minimize Equation (15). We experimentally set λ_{GB} and λ_{FTV} to 0.2 and 0.1, respectively.

4. Experiments

In this section, we evaluate the performance of the existing baselines and our proposed DGGNet on both synthetic and realistic speckle noise reduction tasks.

We compare our method with several popular despeckling methods (including OBNLM [23], SRAD [8], NLLRF [7], Multiscale Hybrid Model (MHM) [35]), the representative deep learning natural image denoising methods (including DnCNN [16], RIDNet [17]), and the latest ultrasound image denoising method (i.e., MSANN [20]). We use two widely used metrics,

Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) to quantify the image quality. We also conduct ablation studies to analyze the effectiveness of different components of our DGGNet.

4.1. Dataset and Training Settings

In this work, we train the neural networks with two datasets. The first dataset is synthetic speckle-degraded images, which are generated according to the multiplicative speckle noise model in [8]. It simulates a noisy image by convolving a 2D point spread function with the ground truth image using a Radio-Frequency (RF) image. We use 5080 patches of size 128 × 128 from the COCO dataset [36] as the noise-free images.

The second dataset used in this study is sourced from Saset INSIGHT 37C (Saset Healthcare Inc., Cheng Du, China), which captures real ultrasound images with speckle noise. During the data collection process, all optimized functions on the machine were turned off to obtain raw ultrasound images with significant speckle noise. More than 3000 noisy images of the liver were collected. The noisy ultrasound images were then processed by the noise reduction function of the INSIGHT 37C to produce images that approximate real noise-free ultrasound images. These processed images are considered ground truth. All noisy and ground truth image pairs were split into 128 × 128 image patches. After this initial processing, experienced doctors were invited to review and select suitable noisy and ground truth image pairs for inclusion in the final ultrasound image dataset. In our experiment, 5040 image pairs were used for training, 1600 for validation, and 3760 for the final test.

The proposed DGGNet is implemented on the PyTorch framework and a computer with four Nvidia GeForce RTX 3090 GPUs. This network is trained for 100 epochs, and the batch size is 8. The Adam algorithm [37] is applied to optimize the proposed DGGNet. The Warmup Scheduler [38] was used in the first three epochs, and then CosineAnnealingLR [39] was utilized for the learning rate between 1×10^{-4} to 1×10^{-6} in the subsequent training epochs. Peak Signal-to-Noise Ratio (PSNR) [40] and Structural Similarity Index (SSIM) [41] are used as the metrics in this work.

4.2. Experimental Results on Synthetic Noisy Images and the Ultrasound Image Dataset

After training on the COCO dataset [36], we opted to validate our model on these five popular datasets (i.e., McMaster [42], Kodak24 [43], BSD68 [44], Set12 [16], and Urban100 [45]). We generated synthetic noisy images by adding speckle noise with the parameters $f_0 \in [5.0, 10.0]$ MHz, $\sigma_x = 0.1$ µs, and $\sigma_y = 0.15$ ms. We also used real speckle noise data captured from Saset INSIGHT 37C.

(I) ***Quantitative Analysis:*** Table 1 shows the PSNR and SSIM values for each method on these datasets. Our DGGNet method can outperform the baselines in most cases. This demonstrates the strong generalization ability of our DGGNet method in producing high-quality despeckled images.

Table 1. PSNR and SSIM results from various despeckling methods on the McMaster [42], Kodak24 [43], BSD68 [44], Set12 [16], and Urban100 [45] datasets for speckle noise. The best results are emphasized in bold.

	Mcmaster		**Kodak24**		**BSD68**		**set12**		**URBAN100**	
Methods	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
SRAD [8]	23.69	0.773	22.55	0.682	22.88	0.744	21.74	0.743	20.71	0.643
OBNLM [23]	23.66	0.623	22.37	0.582	22.90	0.723	22.37	0.724	21.69	0.703
NLLRF [7]	22.81	0.583	21.52	0.498	22.82	0.674	21.50	0.672	20.92	0.631
MHM [35]	29.56	0.856	28.42	0.801	28.22	0.827	28.56	0.838	27.33	0.827
DnCNN [16]	28.30	0.81	27.35	0.76	26.25	0.79	24.54	0.80	25.15	0.75
RIDNet [17]	31.42	0.853	30.11	0.798	**30.20**	0.832	29.14	0.833	27.0	0.81
MSANN [20]	29.53	0.851	28.02	0.791	28.11	0.823	28.53	0.835	27.21	0.825
DGGNet	**31.52**	**0.863**	**30.15**	**0.805**	30.17	**0.832**	**29.15**	**0.845**	**28.23**	**0.835**

For quantitative results of the real ultrasound dataset, we trained our DGGNet on the training set and test it on the validation set. Table 2 shows our established ultrasound speckle noise dataset's quantitative results (PSNR/SSIM). It shows that our DGGNet achieved the best PSNR and SSIM values on this dataset, demonstrating its robustness and effectiveness. Table 2 supports the claim that the deep learning method can obtain better results in the denoising task.

Table 2. The quantitative results on our ultrasound speckle noise dataset. The best results are emphasized in bold.

Method	SRAD [8]	OBNLM [23]	NLLRF [7]	MHM [35]	DnCNN [16]	RIDNet [17]	MSANN [20]	DGGNet
PSNR	26.90	27.23	27.70	28.30	30.37	33.78	33.80	**34.38**
SSIM	0.682	0.745	0.794	0.765	0.899	0.923	0.852	**0.929**

(II) ***Visual Analysis:*** Figure 4 compares the visual effects of different despeckling methods. SRAD [23] and OBNLM [8] smooth the edges and textures too much, losing fine details. NLLRF [7] does not remove enough noise, resulting in a low-quality image. MHM [35] and DnCNN [16] reduce noise well, but they tend to over-smooth the denoised images, especially in regions with fine details or textures. RIDNet [17] over-smooths some fine details and textures in the denoised images. MSANN [20] can improve the denoising results but suffers under or over-smooths in some regions. Our DGGNet method can reduce noise while preserving the integrity of edges, textures, and fine details.

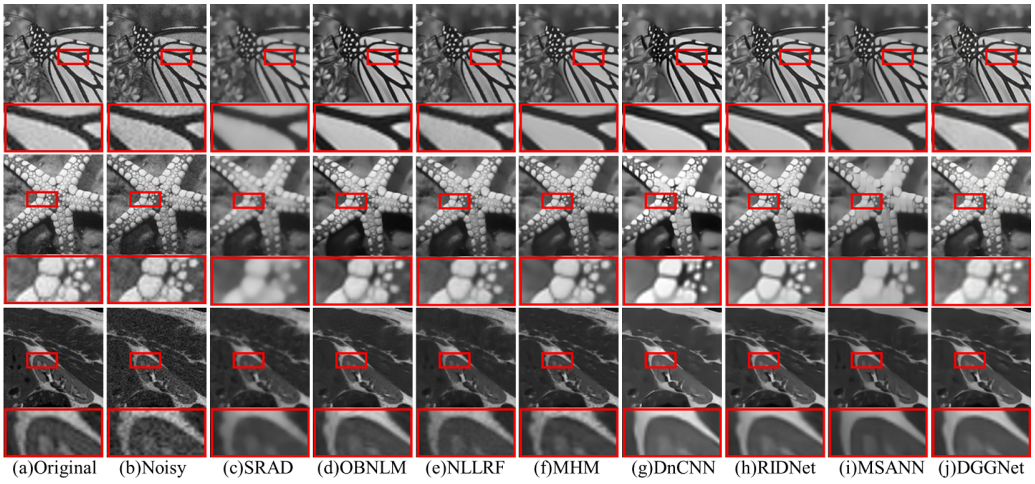

Figure 4. Denoising visualization of our proposed DGGNet comparing competing methods on the ultrasound dataset. From left to right, we show the clean, noisy, and denoising results of SRAD [23], OBNLM [8], NLLRF [7], MHM [35], DnCNN [16], RIDNet [17], MSANN [20] and our proposed DGGNet.

In this section, we also visualize the denoised results of our proposed method with those competing methods on a challenging image in our ultrasound image dataset. As shown in Figure 5, it is clear that our proposed method can restore ultrasound images and preserve the structure better. However, SRAD [23], OBNLM [8], and NLLRF [7] are over-smooth images, and NLLRF [7], DnCNN [16], RIDNet [17], and MSANN [20] cannot restore image structure well. A close inspection of Figure 5 shows that our proposed method can restore the noisy ultrasound image to the ground truth with better preservation structure and does not over-smooth. The visual comparisons show that our DGGNet can effectively suppress the speckle noise while preserving the fine details and textures, whereas the other methods tend to produce over-smoothed or unnatural results.

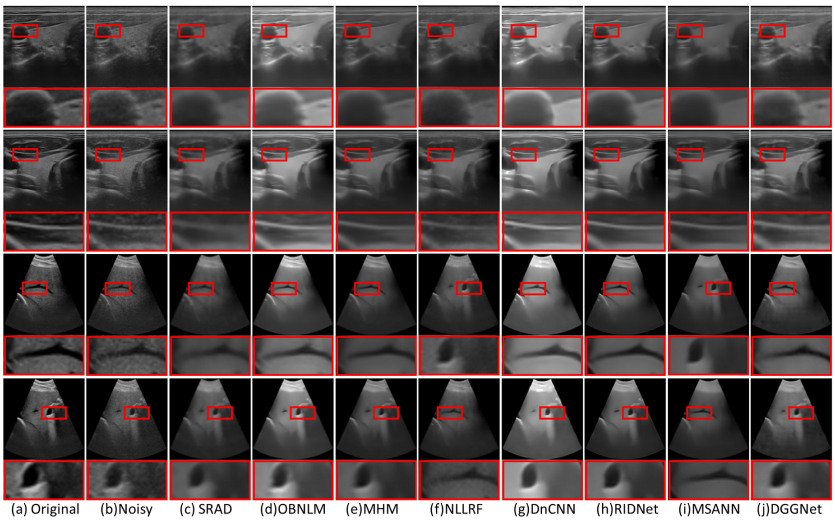

Figure 5. Denoising visualization of our proposed DGGNet comparing competing methods on the ultrasound dataset. From left to right, we show the ground truth, noisy, and denoising results of SRAD [23], OBNLM [8], NLLRF [7], DnCNN [16], MHM [35], RIDNet [17], MSANN [20], and our DGGNet.

4.3. Realistic Experiments on Various Types of Images

Our DGGNet can effectively reduce speckle noise in various imaging systems, such as radar imaging, medical ultrasound imaging, and underwater sonar imaging. We provide the performance of our method on realistic imaging scenarios in Figure 6. This figure shows that our method can preserve more image details and features than traditional methods by addressing the lack of adaptability to speckle noise in traditional approaches. OBNLM [8] and MHM [35] blur the edges and textures of the regions of interest. NLLRF [7] and SRAD [23] do not remove enough noise, resulting in a low-quality image. DnCNN [16], RIDNet [17], and MSANN [20] had strong learning abilities and produced enhanced images with better visual appearance. However, they still had some problems with over-smoothing or insufficient noise removal. Thanks to the use of gradient branch models that enhance the network's ability to preserve edges and the fractional total variation regularization's ability preserve texture, the DGGNet outperforms other methods in terms of visual quality.

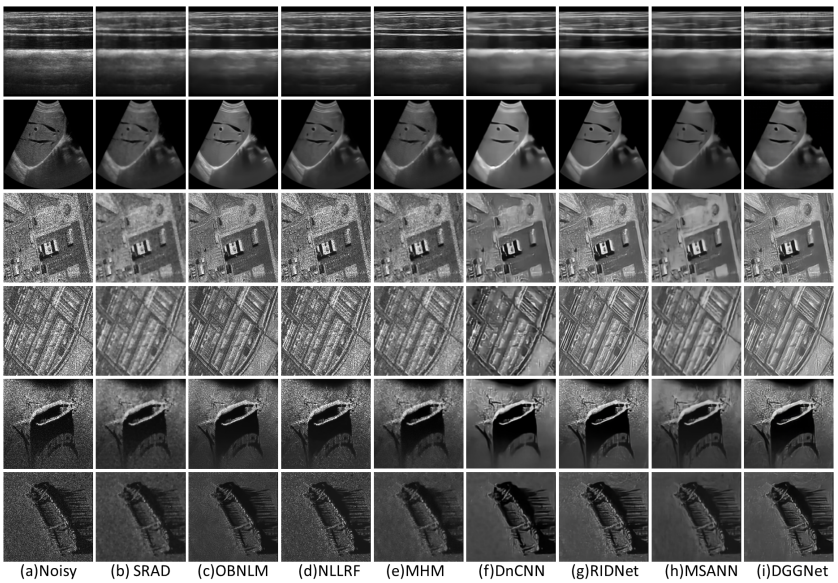

Figure 6. Denoising visualization of our proposed DGGNet compares competing methods on the realistic experiments data. From left to right, we show the noisy, denoising results of SRAD [23], OBNLM [8], NLLRF [7], MHM [35], DnCNN [16], RIDNet [17], MSANN [20] and our proposed DGGNet.

4.4. Ablation Study

The ablation study is performed on the provided real speckle noise dataset. The backbone network of our proposed DGGNet is Unet [46]. Our model consists of four special function blocks: gradient branch architecture, context block, self-attention block, and fractional total variation. The results of the ablation study are shown in Table 3. When all components are available, the proposed method achieves the best performance, while the absence of any of the above components will degrade the performance. The ablation study shows that all components of our DGGNet contribute to the final denoising result.

Table 3. Ablation study of different components. Both PSNR and SSIM are tested on our real speckle noise ultrasound image. The best results are emphasized in bold.

Unet	✓	✓	✓	✓	✓	✓	✓	✓
Context block	×	×	✓	×	×	✓	✓	✓
gradient branch	×	✓	×	×	✓	×	✓	✓
Self-attention block	×	×	×	✓	✓	✓	✓	✓
fractional total variation	×	×	×	×	×	×	×	✓
PSNR	33.16	33.27	33.43	33.81	33.83	34.03	34.09	**34.38**
SSIM	0.919	0.920	0.922	0.923	0.924	0.927	0.928	**0.929**

As discussed above, our network utilizes the gradient branch to guide the denoising branch for preserving image structure. The denoising branch obtains additional multi-level structure features from the gradient branch. To demonstrate the effectiveness of the gradient branch, we visualize the mean feature map of the image denoising decoding architecture in Figure 7. The visualization result shows that in the gradient regions, the map feature that is obtained by the denoising network with the guidance of the gradient branch can appear more distribution of highlight values. However, the region without the gradient branch is blurry and has lower brightness. Both Table 3 and Figure 7 support our claim that the denoising network can perform better in the sharp regions' recovery and structure preservation with the guide of the gradient branch.

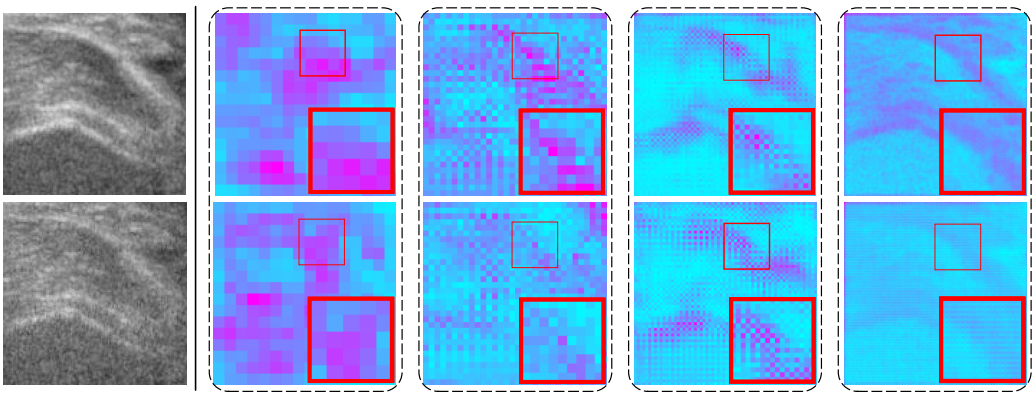

Figure 7. Average feature maps of results of the upsampling block in the decoding architecture of the denoising branch in our proposed DGGNet. The top image in (**a**) is our denoising result, and the bottom image is the corresponding noisy image. (**b**–**e**) are the average feature maps of 16×16, 32×32, 64×64, and 128×128 in the denoising branch of the decoding structure. The upper images of those image pairs are the average feature map of the denoising branch with the gradient branch, while the lower images are not. This shows that with the guide of the gradient branch in our DGGNet, the denoising result can preserve structure information better.

5. Conclusions

In this paper, we propose a structure-preserving coherent image-denoising network, termed DGGNet, which aims to effectively balance image denoising and structural preservation. Our network employs a gradient branch to guide the denoising process and mitigate

the problem of over-smoothing. Additionally, we introduce a fractional total variational loss function to avoid detail ambiguity and reduce the excessive smoothing often associated with integral-order methods.

The quantitative results, as shown in Table 2, highlight that DGGNet significantly outperforms traditional methods in terms of PSNR and SSIM on the ultrasound speckle noise dataset. Our method achieves a PSNR of **34.38** and an SSIM of **0.929**, surpassing other methods such as SRAD, OBNLM, and NLLRF. Table 1 further confirms that DGGNet provides the highest PSNR and SSIM values across various datasets, including McMaster, Kodak24, BSD68, Set12, and Urban100, demonstrating its superior performance compared to DnCNN, RIDNet, and MSANN.

In summary, DGGNet represents a significant advancement in denoising speckle noise, providing improved quality for medical imaging, underwater sonar, and synthetic aperture radar applications. We believe that our method's integration of cutting-edge preprocessing techniques and advanced network design offers a substantial improvement over existing deep learning-based approaches.

6. Future Work

In future work, we aim to enhance denoising techniques by exploring advanced regularization strategies, hybrid methods, adaptive approaches, and extensive real-world testing to improve both noise suppression and structural detail preservation. Additionally, we plan to integrate our DGGNet with other models, including generative models, ensemble techniques, cross-modal approaches, and advanced deep learning frameworks, to leverage their complementary strengths and achieve superior performance. These efforts will help advance the state-of-the-art in coherent image denoising and extend the applicability of our methods across diverse imaging domains.

Author Contributions: Conceptualization, H.Y.; methodology, L.W.; software, L.W.; validation, L.W. and Y.-F.P.; formal analysis, H.Y.; investigation, P.L.; resources, P.L.; data curation, P.L.; writing—original draft preparation, L.W. and J.L.; writing—review and editing, Y.-F.P.; visualization, J.L.; supervision, Y.-F.P.; project administration, L.W.; funding acquisition, Y.-F.P. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the National Natural Science Foundation of China under Grants 62171303, China South Industries Group Corporation (Chengdu) Fire Control Technology Center Project (non-secret) under Grants HK20-03, and the National Key Research and Development Program Foundation of China under Grants 2018YFC0830300.

Institutional Review Board Statement: Ethical review and approval were waived for this study because the data used in the research were obtained from two sources: (1) publicly available datasets, which do not contain any personally identifiable information and are widely recognized as ethical for use in research, and (2) data collected from volunteers who provided informed consent prior to participation. All volunteer data were gathered in compliance with legal and ethical standards, ensuring that participants were fully aware of the nature and purpose of the study, as well as their rights to withdraw at any time. Therefore, this study did not require further ethical approval.

Informed Consent Statement: Informed consent was obtained from all volunteers involved in the study. Prior to participation, each volunteer was fully informed about the nature, purpose, and procedures of the research. They were made aware of their right to withdraw from the study at any time without any consequences. The data collected were anonymized to protect the privacy of the participants, and all procedures were conducted in compliance with relevant legal and ethical standards.

Data Availability Statement: The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: We express our gratitude to Saset Healthcare Inc. for generously providing the authentic ultrasound image data and extend our appreciation to engineers Jueli Chen and Youling Shi for their invaluable assistance in collecting these real ultrasound images. We express our gratitude to all anonymous reviewers for their valuable comments on this article.

Conflicts of Interest: All authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
DGGNet	Deep Gradient-Guidance Networks
DnCNN	Denoising Convolutional Neural Network
DRAN	Deep Residual Attention Network
MARU	Mixed-Attention Residual U-Net
MHM	Multiscale Hybrid Model
MSANN	Multi-Scale Attention Network
NLLRF	Non-Local Low-Rank Filter
NLM	Non-Local Means
OBNLM	Order-Based Nonlocal Means
PSNR	Peak Signal-to-Noise Ratio
RF	Radio-Frequency
RIDNet	Residual Image Denoising Network
SAR	Synthetic Aperture Radar
SRAD	Speckle Reducing Anisotropic Diffusion
SSIM	Structural Similarity Index
U-Net	U-Net Network

References

1. Wang, X.; Jiao, J.; Yin, J.; Zhao, W.; Han, X.; Sun, B. Underwater sonar image classification using adaptive weights convolutional neural network. *Appl. Acoust.* **2019**, *146*, 145–154. [CrossRef]
2. Iwin Thanakumar Joseph, S.; Sasikala, J.; Sujitha Juliet, D.; Velliangiri, S. Hybrid spatio-frequency domain global thresholding filter (HSFGTF) model for SAR image enhancement. *Pattern Recognit. Lett.* **2021**, *146*, 8–14. [CrossRef]
3. Mugasa, H.; Dua, S.; Koh, J.E.; Hagiwara, Y.; Lih, O.S.; Madla, C.; Kongmebhol, P.; Ng, K.H.; Acharya, U.R. An adaptive feature extraction model for classification of thyroid lesions in ultrasound images. *Pattern Recognit. Lett.* **2020**, *131*, 463–473. [CrossRef]
4. Sagheer, S.V.M.; George, S.N. A review on medical image denoising algorithms. *Biomed. Signal Process. Control* **2020**, *61*, 102036.
5. Baselice, F.; Ferraioli, G.; Pascazio, V.; Schirinzi, G. Enhanced Wiener Filter for Ultrasound image denoising. In Proceedings of the European Medical and Biological Engineering Conference Nordic-Baltic Conference on Biomedical Engineering and Medical Physics, Tampere, Finland, 11–15 June 2017; pp. 65–68. [CrossRef]
6. Loupas, T.; McDicken, W.; Allan, P.L. An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *IEEE Trans. Circuits Syst.* **1989**, *36*, 129–135. [CrossRef]
7. Zhu, L.; Fu, C.W.; Brown, M.S.; Heng, P. A Non-local Low-Rank Framework for Ultrasound Speckle Reduction. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 493–501.
8. Yu, Y.; Acton, S. Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.* **2002**, *11*, 1260–1270.
9. Vazquez-Corral, J.; Bertalmío, M. Angular-Based Preprocessing for Image Denoising. *IEEE Signal Process. Lett.* **2018**, *25*, 219–223. [CrossRef]
10. Shen, Y.; Liu, Q.; Lou, S.; Hou, Y. Wavelet-Based Total Variation and Nonlocal Similarity Model for Image Denoising. *IEEE Signal Process. Lett.* **2017**, *24*, 877–881. [CrossRef]
11. Zhou, Y.; Li, Y.; Guo, Z.; Wu, B.; Zhang, D. Multiplicative Noise Removal and Contrast Enhancement for SAR Images Based on a Total Fractional-Order Variation Model. *Fractal Fract.* **2023**, *7*, 329. [CrossRef]
12. Bi, S.; Li, M.; Cai, G. Mixed Fractional-Order and High-Order Adaptive Image Denoising Algorithm Based on Weight Selection Function. *Fractal Fract.* **2023**, *7*, 566. [CrossRef]
13. Song, Y.; Zhu, Y.; Du, X. Grouped Multi-Scale Network for Real-World Image Denoising. *IEEE Signal Process. Lett.* **2020**, *27*, 2124–2128. [CrossRef]
14. Bai, Y.C.; Zhang, S.; Chen, M.; Pu, Y.F.; Zhou, J.L. A Fractional Total Variational CNN Approach for SAR Image Despeckling. In *Intelligent Computing Methodologies, Proceedings of the 14th International Conference, ICIC 2018, Wuhan, China, 15–18 August 2018*; Proceedings, Part III 14; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; pp. 431–442.
15. Liang, J.; Liu, R. Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network. In Proceedings of the 2015 8th International Congress on Image and Signal Processing (CISP), Shenyang, China, 14–16 October 2015; pp. 697–701.
16. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [CrossRef] [PubMed]

17. Anwar, S.; Barnes, N. Real Image Denoising With Feature Attention. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 3155–3164.
18. Sharif, S.; Naqvi, R.A.; Biswas, M. Learning medical image denoising with deep dynamic residual attention network. *Mathematics* **2020**, *8*, 2192. [CrossRef]
19. Lan, Y.; Zhang, X. Real-Time Ultrasound Image Despeckling Using Mixed-Attention Mechanism Based Residual UNet. *IEEE Access* **2020**, *8*, 195327–195340. [CrossRef]
20. Guo, Y.; Lu, Y.; Liu, R.W.; Zhu, F. Blind Image Despeckling Using Multi-Scale Attention-Guided Neural Network. *IEEE Trans. Artif. Intell.* **2023**, *5*, 205–216. [CrossRef]
21. Pu, Y.F. Fractional-Order Euler-Lagrange Equation for Fractional-Order Variational Method: A Necessary Condition for Fractional-Order Fixed Boundary Optimization Problems in Signal Processing and Image Processing. *IEEE Access* **2016**, *4*, 10. [CrossRef]
22. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 21–23 September 2005; Volume 2, pp. 60–65. [CrossRef]
23. Coupe, P.; Hellier, P.; Kervrann, C.; Barillot, C. Nonlocal Means-Based Speckle Filtering for Ultrasound Images. *IEEE Trans. Image Process.* **2009**, *18*, 2221–2229. [CrossRef]
24. Jia, X.; Feng, X.; Wang, W. Rank constrained nuclear norm minimization with application to image denoising. *Signal Process.* **2016**, *129*, 1–11. [CrossRef]
25. Tirer, T.; Giryes, R. Super-Resolution via Image-Adapted Denoising CNNs: Incorporating External and Internal Learning. *IEEE Signal Process. Lett.* **2019**, *26*, 1080–1084. [CrossRef]
26. Wang, Y.; Song, X.; Chen, K. Channel and Space Attention Neural Network for Image Denoising. *IEEE Signal Process. Lett.* **2021**, *28*, 424–428. [CrossRef]
27. He, K.; Sun, J.; Tang, X. Guided Image Filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1397–1409. [CrossRef] [PubMed]
28. Duman, E.A. An Edge Preserving Image Denoising Framework Based on Statistical Edge Detection and Bilateral Filter. *Mehmet Akif Ersoy Üniversitesi Fen Bilim. EnstitüSü Derg.* **2021**, *12*, 519–531. [CrossRef]
29. Bai, J.; Feng, X.C. Fractional-Order Anisotropic Diffusion for Image Denoising. *IEEE Trans. Image Process.* **2007**, *16*, 2492–2502. [CrossRef] [PubMed]
30. Pu, Y.F.; Zhou, J.L.; Yuan, X. Fractional differential mask: A fractional differential-based approach for multiscale texture enhancement. *IEEE Trans. Image Process.* **2010**, *19*, 491–511.
31. Lysaker, M.; Osher, S.; Tai, X.C. Noise removal using smoothed normals and surface fitting. *IEEE Trans. Image Process.* **2004**, *13*, 1345–1357. [CrossRef]
32. Chen, D.; Sun, S.; Zhang, C.; Chen, Y.; Xue, D. Fractional-order TV-L2 model for image denoising. *Cent. Eur. J. Phys.* **2013**, *11*, 1414–1422. [CrossRef]
33. Podlubny, I.; Chechkin, A.; Skovranek, T.; Chen, Y.Q.; Jara, B.M.V. Matrix approach to discrete fractional calculus II: Partial fractional differential equations. *J. Comput. Phys.* **2008**, *228*, 3137–3153. [CrossRef]
34. Kanopoulos, N.; Vasanthavada, N. Design of an image edge detection filter using the Sobel operator. *IEEE J. Solid-State Circuits* **1988**, *23*, 358–367. [CrossRef]
35. Wang, L.; Pu, Y.F.; Liu, P.; Hao, Y. Multiscale hybrid method for speckle reduction of medical ultrasound images. *Multimed. Tools Appl.* **2023**, *83*, 55219–55234. [CrossRef]
36. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollar, P.; Zitnick, L. Microsoft COCO: Common Objects in Context. In Proceedings of the European Conference on Computer Vision, ECCV, Zurich, Switzerland, 6–12 September 2014.
37. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
38. Xiong, R.; Yang, Y.; He, D.; Zheng, K.; Zheng, S.; Xing, C.; Zhang, H.; Lan, Y.; Wang, L.; Liu, T. On Layer Normalization in the Transformer Architecture. In Proceedings of the 37th International Conference on Machine Learning, PMLR, Vienna, Austria, 12–18 July 2020.
39. Wang, Y.; Solomon, J. Deep Closest Point: Learning Representations for Point Cloud Registration. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 3522–3531.
40. Zhang, J.; Lin, G.; Wu, L.; Wang, C.; Cheng, Y. Wavelet and fast bilateral filter based de-speckling method for medical ultrasound images. *Biomed. Signal Process. Control* **2015**, *18*, 1–10. [CrossRef]
41. Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
42. Zhang, L.; Wu, X.; Buades, A.; Li, X. Color demosaicking by local directional interpolation and nonlocal adaptive thresholding. *J. Electron. Imaging* **2011**, *20*, 023016.
43. Franzen, R. Kodak Lossless True Color Image Suite. Available online: http://r0k.us/graphics/kodak (accessed on 12 September 2023).
44. Roth, S.; Black, M. Fields of Experts: A framework for learning image priors. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 21–23 September 2005; Volume 2, pp. 860–867.

45. Huang, J.B.; Singh, A.; Ahuja, N. Single image super-resolution from transformed self-exemplars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5197–5206.
46. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015*; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.