Attention-guided U-Net Model with Improved Residual Blocks for Ultrasound Image Denoising

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

Image denoising is one of the fundamental challenges in medical imaging which enhances image quality by suppressing noise from a noisy image. Ultrasound imaging as one of the most powerful tools for clinical diagnosis, may not be properly utilized due to the presence of speckle noise. Although many of the models designed for despeckling of ultrasound images have been originally developed for segmentation tasks, they still exhibit remarkable success and accuracy in the despeckling domain. In this study, we investigate the recent despeckling performance of U-Net based networks and propose a novel approach, called Attention guided with Improved Residuals U-Net (AIR U-Net), which incorporates improved residual blocks and an attention mechanism in a U-Net architecture. Our experimental evaluation with the DND dataset demonstrates superior performance compared to a number of existing ultrasound denoising algorithms.Our code is publicly available at https://github.com/mselmangokmen/udea.

Keywords: Deep learning, Image denoising, speckle noise, U-Net, Ultrasound

1. Introduction

Medical images such as ultrasound images are often affected by noise during acquisition, processing, and transmission. In the field of medical imaging, it is common practice to use various techniques such as smoothing and thresholding to recover the original image. An image denoising algorithm takes a noisy medical image as input and outputs a denoised image, which is essential for accurate diagnosis and treatment planning. The purpose of denoising in medical imaging extends beyond improving the visual quality of images. Denoising is a critical step in the pipeline for higher-level tasks such as image segmentation, classification, and object recognition, which play an important role in accurate diagnosis and treatment (Fu et al., 2020).

U-Net models are commonly used for segmentation and detection tasks in computer vision, while U-Net based models for image denoising are relatively limited (Bian et al., 2020). In recent years, U-Net based models have shown remarkable performance in various computer vision applications, motivating researchers to explore the use of U-Net for image denoising. In this context, several recent studies have investigated the effectiveness of U-Net models for denoising tasks (Azad et al., 2022). Motivated by the success of U-Net based models in various computer vision applications, there is a growing interest in exploring their potential for image denoising. In this context, this paper aims to investigate the effectiveness of U-Net based models for denoising tasks. Specifically, we focus on U-Net models designed for segmentation and detection tasks, and explore their performance on denoising tasks. Our study will provide insights into the potential of U-Net models for image denoising, and contribute to the development of more effective denoising methods.

2. Related Works

Until the present time, numerous techniques have been proposed for despeckling ultrasound images.Generally, the existing despeckling methods fall into two categories (Fan et al., 2019): frequency domain-based and spatial domain-based methods. Among frequency domain-based methods, wavelet-based methods (Gan et al., 2019) are quite popular. These methods function by converting speckle noise to additive noise, then eliminating it within the wavelet domain. However, the efficacy of these methods for despeckling is hampered since speckle noise in actual ultrasound images is not purely multiplicative noise, and the selection of mother wavelet may introduce artifacts.

Traditional spatial domain based methods use local pixel comparison, but they are not effective in reducing noise while preserving image details. Instead of individual pixels, the NLM method (Buades et al., 2005) explores the self-similarities between image patches to restore each pixel with the weighted average of all pixels in a search window. However, this method is not suitable for speckle noise reduction. To overcome this limitation, modified NLM approaches have been proposed, such as modified BM3D algorithm (Burger et al., 2012). These methods use various techniques, such as pre-filtering with local noise statistics or using different patch distance measures.

The conventional despeckling methods mentioned earlier are limited in their ability to reduce noise while preserving image details, especially in cases of high speckle corruption. Moreover, these methods are often not capable of real-time ultrasound image despeckling due to their complex operations. On the other hand, deep learning methods offers a promising solution for real-time and effective ultrasound image despeckling. As a popular algorithm in the field of machine learning, it can automatically learn intrinsic features from training data and enable highly efficient image denoising.

Various deep learning models, such as auto-encoders, CNNs, and RNNs, have been proposed for ultrasound image despeckling. CNNs are the most popular and successful in image processing and analysis tasks, including denoising (Tian et al., 2019). While the DCNND and DnCNN models perform well in removing Gaussian noise using residual learning, they cannot accurately estimate the residual term for speckle noise removal (Bian et al., 2020).

On the other hand, U-Net models have gained popularity in image denoising due to their superior performance in preserving fine image details and texture. They have been shown to outperform traditional denoising methods such as BM3D and wavelet-based methods in terms of both quantitative metrics and visual quality of the denoised images. U-Net models have become a popular choice for image denoising tasks in various applications, including medical imaging and computer vision. Studies have demonstrated that U-Net models not only outperform other denoising methods in terms of PSNR and SSIM but also preserve important image features and edges more effectively (Jia et al., 2021). In this context, this paper aims to investigate the effectiveness of U-Net based models for denoising tasks.Specifically, we focus on U-Net models designed for segmentation and detection tasks, and explore their performance on denoising tasks.

Our study will provide insights into the potential of U-Net models for image denoising, and contribute to the development of more effective denoising methods. First model we will investigate is ResU-Net (Zhang et al., 2017) mostly used for denoising and segmentation tasks and contributes a novel approach in literature. Following one is Attention U-Net

(Pires et al., 2021), especially developed for pancreas segmentation. Despite Attention U-Net improved for segmentation tasks, the attention gates are provided in the model shows superior performance on denoising tasks.The last model (Zhang et al., 2022), we will evaluate is developed for ultrasound image denoising tasks and that will aid us for comparing our model's success in same research area.

3. Methodology

3.1. Data Registration

Proposed model Air U-Net takes the advantage of small patch-based images generated from DND dataset (Plötz and Roth, 2017) with the size of 64x64. Small image patches don't carry global features comprehensively and that aids the model focus on pixel-wise noise detection while training instead of focusing global features. The reason for using small patch-based images is to avoid global features that are spread throughout high-resolution noisy images.The speckle distribution of a small patch-based image with dimensions of 64x64 is shown in figure 1 . The image derivatives obtained from the subtraction of the noisy and clean images are also depicted. It is explicitly clear that noise distribution is located around tissue change.

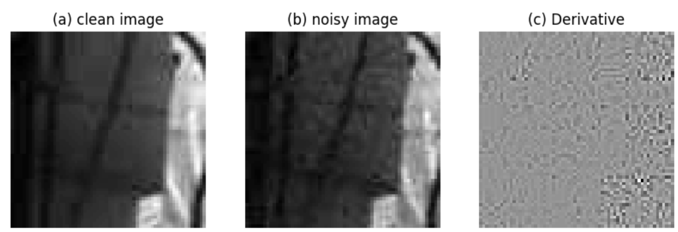


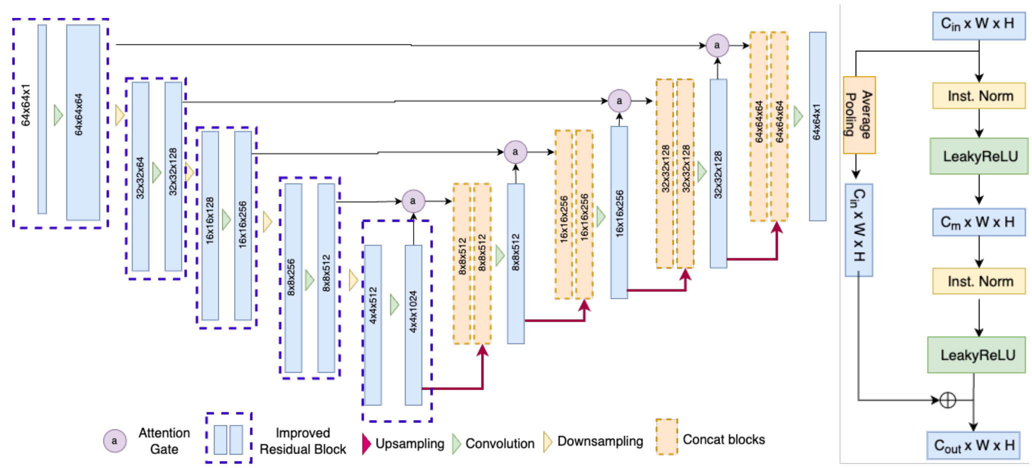
Figure 1: Representation of noise distribution on real-life images.

3.2. Adding Speckle Noise

In our study, we added speckle noise to ultrasound images using the Rayleigh distribution, which is commonly used to model speckle noise in ultrasound imaging. During the training process, we utilized four different Rayleigh noise levels, namely 0.1, 0.25, 0.5, and 0.75, to evaluate the effectiveness of our proposed denoising techniques under varying levels of noise. The speckle noise added to the images accurately mimics the type of noise commonly found in ultrasound images, and allows us to test the robustness of our denoising model in real-world scenarios.

3.3. Proposed Model

The improved residual blocks are utilized in this study, which executes an average pooling layer in addition to traditional residual blocks (Deng, 2021). The average pooling layer acts as a simple form of noise reduction by smoothing out the input signal and reducing the influence of individual noisy pixels. Moreover, the attention gates (Pires et al., 2021) used between each of the encoder and decoder layers, can selectively enhance informative features while suppressing irrelevant ones. This enables the model to better distinguish signal from noise, leading to improved performance in denoising medical images.



(a) AIR U-Net

(b) Improved residual block

Figure 2: Proposed AIR U-Net model takes a noisy ultrasound frame as input and outputs the clean image.

In this study, we compared the denoising performance of our proposed model with three other models using the DND dataset, which contains real-life noisy images captured by different devices with low and high resolutions (Plötz and Roth, 2017). The use of this dataset was primarily to leverage the absence of Rayleigh distribution noise and enhance the denoising performance by improving the interpretation of noise distribution in real-life images. For testing, we employed the US-4 dataset, which includes ultrasound video frames (Chen et al., 2020). Ultrasound videos have the potential to reveal more noisy frames during diagnosis compared to static images created for segmentation and other tasks.

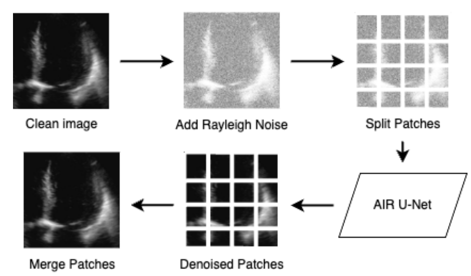
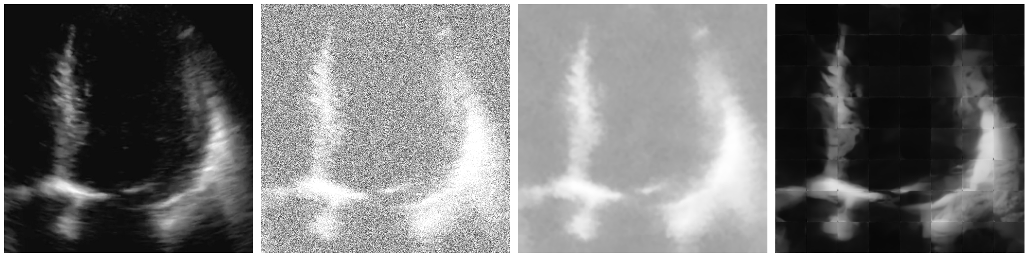


Figure 3: Illustration of the AIR U-Net based ultrasound image denoising procedure.

Table 1: Accuracy metrics during training in four different noise levels

Methods	SSIM			
	$\sigma = 0.10$	$\sigma = 0.25$	$\sigma = 0.50$	$\sigma = 0.75$
BM3D (Burger et al., 2012)	65.5%	40.1%	23.9%s	18.2%
ResU-Net (Zhang et al., 2017)	94.6%	92.0%	88.2%	86.3%
Attention U-Net (Pires et al., 2021)	92.1%	89.3%	86.7%	84.2%
RatU-Net (Zhang et al., 2022)	94.7%	89.4%	89.0%	82.1%
AIR U-Net (Ours)	**96.1%**	**93.5%**	**90.3%**	**88.4%**



(a) Clean image (b) Noisy image (c) BM3D (d) AIR U-Net

Figure 4: Visual comparison suggests superior denoising of a noisy image by our AIR U-Net over the baseline BM3D model.

4. Experiment Setup

During our experiments, optimum training parameters evaluated and also best validation results recorded for each U-Net based model.

4.1. Dataset

During the training process, we constructed a training dataset using the DND dataset, which consists of 6144 small images with a resolution of 64x64. These smaller images were generated from 96 noisy images with a resolution of 512x512. The Figure 3 represents running progress for data registration.

4.2. Implementation

The performance of the Air U-Net model was evaluated against four baseline models. In order to thoroughly assess its performance, the Air U-Net model was trained for each of the four Rayleigh noise levels (0.1, 0.25, 0.5, and 0.75) separately, for a total of 800 epochs using the Adam optimizer. The other models used in the training process (ResU-Net, Attention U-Net, and Rat-Unet) were also trained for 200 epochs using the same optimizer. To achieve optimal training results, the learning rate was adjusted from $1e{-2}$ to $1e{-5}$, and the batch size was changed from 200 to 50 based on the training results.
Table 1 reports the denoising performance for three different U-Net based architectures. These models were trained for 200 epochs using four different Rayleigh noise distributions. All the models were saved separately for every noise level. This resulted in a total of 800 epochs for training each U-Net based network. Our experiments on the DND dataset (Plötz and Roth, 2017) have shown that the noise distribution is mainly concentrated around the edges of tissue changes in real-life images. Based on this observation, we separated the images into 64×64 patches during the training and testing stages, and focused on the distribution of noise rather than global features. Using the DND dataset in our training and testing stages was essential for interpreting the natural noise in images. Our AIR U-Net demonstrated superior performance in ultrasound image denoising, as it is capable of learning the characteristics of noise using 64×64 patch images during training. This approach increases the model's attention from global features to the noise distribution in small patches, resulting in better noise extraction compared to larger 512×512 images.

5. Tests and Results

5.1. Dataset

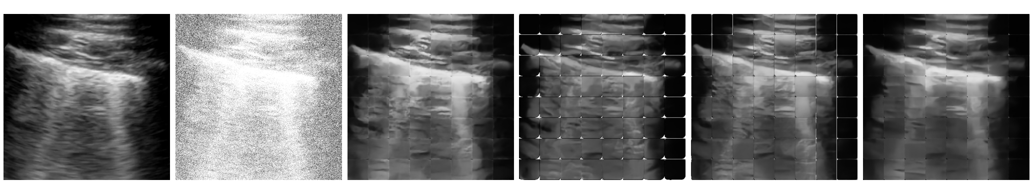
To test the performance of the proposed model in real-time medical imaging, we selected the US-4 dataset which contains four different convex probes and two scan regions (i.e., lung and liver). The images in the US-4 dataset (Chen et al., 2020) represent each frame in video scans, making it a suitable dataset for testing the performance of our denoising model in a realistic setting.

5.2. Tests

To test the performance of the proposed model in real-time medical imaging, we selected the US-4 dataset which contains four different convex probes and two scan regions (i.e., lung and liver). The images in the US-4 dataset represent each frame in video scans, making it a suitable dataset for testing the performance of our denoising model in a realistic setting. Figure 4 shows the result of testing our proposed model on a video frame from (Chen et al., 2020). The noisy image is obtained by applying Rayleigh noise to the clean image, which is then segmented into small 64×64 batches for the denoising process using our proposed model as depicted in figure 3. The denoised image is obtained by combining the predicted clean batches. We also compare the performance of our proposed model with BM3D, one of the most popular image denoising algorithms.

5.3. Results

Our complete test results for 0.5 Rayleigh noise level are represented in figure 5. The first image shows target output and second one represents the noisy image with a Rayleigh distribution of 0.5, while the remaining subimages (c, d, e and f) show the denoised results of ResUNet, Attention U-Net, Rat U-Net, and Air U-Net respectively. Furthermore, Table 2 includes the SSIM scores for each model and Rayleigh distribution level (0.1, 0.25, 0.5, 0.75).



(a) Clean image (b) Noisy image 0.5 (c) Res U-Net (d) Attention U-Net (e) Rat U-Net (f) Air U-Net

Figure 5: Visual comparison for denoising performance of Res U-net, Rat U-Net, Attention U-Net and our proposed model Air U-Net.

Table 2: Accuracy test metrics in four different noise levels

Methods	SSIM			
	$\sigma = 0.10$	$\sigma = 0.25$	$\sigma = 0.50$	$\sigma = 0.75$
ResU-Net (Zhang et al., 2017)	87.8%	79.8%	68.7%	62.3%
Attention U-Net (Pires et al., 2021)	77.1%	67.6%	61.5%	52.6%
RatU-Net (Zhang et al., 2022)	85.0%	70.8%	65.1%	54.4%
AIR U-Net (Ours)	**90.7%**	**82.5%**	**70.9%**	**63.9%**

6. Conclusions

In this paper, the recent models are investigated and denoising performance of these models are evaluated. The addition of improved residual network and mixed-attention mechanism significantly improves the denoising performance of this network. Experiments are deployed on four different U-Net shaped models, it is discovered that implementing denoising algorithms with small patched images have 64×64 avoids the model focusing on unnecessary global features and improves on denoising performance. Furthermore, attention gates are employed on our proposed model increase decoding and construction of image significantly.

The proposed model also employs attention gates to enhance image decoding and construction. However, a drawback of the model is that bounding boxes occur after combining ultrasound images from patches. Future studies will explore the impact of different types of convolution blocks and attention gates to improve the model for real-time denoising of ultrasound images. The ultimate goal is to develop an effective model for real-time ultrasound image denoising.

7. Acknowledgments

I would like to acknowledge Abdullah-Al-Zubaer Imran, Ph.D., for his guidance and support throughout this research project. I am grateful for the opportunity to conduct this research at the University of Kentucky under his supervision. His expertise and encouragement have been invaluable in the completion of this work.

References

Reza Azad, Ehsan Khodapanah Aghdam, Amelie Rauland, Yiwei Jia, Avval, Atlas Haddadi, Afshin Bozorgpour, Sanaz Karimijafarbigloo, Joseph Paul Cohen, Ehsan Adeli, and Dorit Merhof. Medical image segmentation review: The success of U-Net. 2022.

Tong Bian, Yuxuan Yi, Jiale Hu, Yin Zhang, Yide Wang, and Lu Gao. A residual-based deep learning approach for ghost imaging. *Sci. Rep.*, 10(1):12149, July 2020.

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)*. IEEE, 2005.

H C Burger, C J Schuler, and S Harmeling. Image denoising: Can plain neural networks compete with BM3D? In *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, June 2012.

Yixiong Chen, Chunhui Zhang, Li Liu, Cheng Feng, Changfeng Dong, Yongfang Luo, and Xiang Wan. USCL: Pretraining deep ultrasound image diagnosis model through video contrastive representation learning. November 2020.

Hao Deng. Research on u-net improvement by attention and residual block. In *ICMLCA 2021; 2nd International Conference on Machine Learning and Computer Application*, pages 1–5, Dec 2021.

Linwei Fan, Fan Zhang, Hui Fan, and Caiming Zhang. Brief review of image denoising techniques. *Vis. Comput. Ind. Biomed. Art*, 2(1):7, July 2019.

Yabo Fu, Yang Lei, Tonghe Wang, Walter J Curran, Tian Liu, and Xiaofeng Yang. Deep learning in medical image registration: a review. *Phys. Med. Biol.*, 65(20):20TR01, October 2020.

Zhenhua Gan, Fumin Zou, Nianyin Zeng, Baoping Xiong, Lyuchao Liao, Han Li, Xin Luo, and Min Du. Wavelet denoising algorithm based on NDOA compressed sensing for fluorescence image of microarray. *IEEE Access*, 7:13338–13346, 2019.

Fan Jia, Wing Hong Wong, and Tieyong Zeng. DDUNet: Dense dense U-Net with applications in image denoising. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. IEEE, October 2021.

Rafael G Pires, Daniel F S Santos, Marcos C S Santana, Claudio F G Santos, and Joao P Papa. Image denoising using Attention-Residual convolutional neural networks. January 2021.

Tobias Plötz and Stefan Roth. Benchmarking denoising algorithms with real photographs. July 2017.

Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin. Deep learning on image denoising: An overview. 2019.

Huibin Zhang, Qiusheng Lian, Jianmin Zhao, Yining Wang, Yuchi Yang, and Suqin Feng. RatUNet: residual U-Net based on attention mechanism for image denoising. *PeerJ Comput. Sci.*, 8(e970):e970, May 2022.

Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road extraction by deep residual U-Net. November 2017.