



Contents lists available at ScienceDirect

# Engineering Science and Technology, an International Journal

journal homepage: [www.elsevier.com/locate/jestch](http://www.elsevier.com/locate/jestch)

## Full Length Article

## Removal of speckle noises from ultrasound images using five different deep learning networks

Onur Karaoglu<sup>a</sup>, Hasan Sakir Bilge<sup>b</sup>, İhsan Uluer<sup>a</sup><sup>a</sup> Department of Electrical & Electronics Engineering, Karabuk University, Karabuk, Turkey<sup>b</sup> Department of Electrical & Electronics Engineering, Gazi University, Ankara, Turkey

## ARTICLE INFO

## Article history:

Received 18 August 2020

Revised 27 May 2021

Accepted 16 June 2021

Available online 08 July 2021

## Keywords:

Ultrasound imaging

Deep learning

Speckle noise

Denoising

Image enhancement

## ABSTRACT

Image enhancement methods are applied to medical images to reduce the noise that they contain. There are many academic studies in the literature using classical image enhancement methods. Ultrasound imaging is a medical imaging method that is used for the diagnosis of diseases. In this study, speckle noises with Rayleigh distribution at four different noise levels ( $\sigma = 0.10, 0.25, 0.50, 0.75$ ) are added to ultrasound images of the brachial plexus nerve region. Five different deep learning networks (Dilated Convolution Autoencoder Denoising Network/Di-Conv-AE-Net, Denoising U-Shaped Net/D-U-Net, BatchRenormalization U-Net/Br-U-Net, Generative Adversarial Denoising Network/DGan-Net, and CNN Residual Network/DeRNet) are used for reducing the speckle noises of the ultrasound images. The performances of the deep networks are compared with block-matching and 3D filtering (BM3D), which is one of the most preferred classical image enhancement algorithms; with classical filters including Bilateral, Frost, Kuan, Lee, Mean, and Median Filters; and with deep learning networks including Learning Pixel-Distribution Prior with Wider Convolution for Image Denoising (WIN5-RB), Denoising Prior Driven Deep Neural Network for Image Restoration (DPDNN), and Fingerprint Image Denoising and Inpainting Using M-Net Based Convolutional Neural Networks (FPD-M-Net). Network performance is evaluated according to peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and runtime criteria and the proposed deep learning networks are shown to outperform the other networks.

© 2021 Karabuk University. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

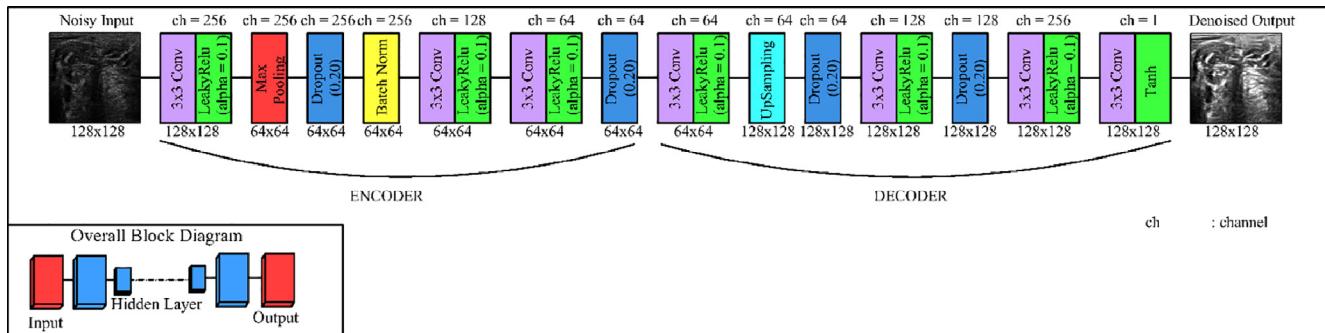
Ultrasound scanners obtain ultrasound images according to the principle of echo imaging. Pulsed acoustic waves generated by the scanner are applied to body tissues [1,2]. Ultrasound imaging is preferred for many reasons. It has many advantages compared to computed tomography and magnetic resonance imaging, such as being more economical, compact, and portable while offering real-time operation and being free of ionizing radiation [2]. However, ultrasound images contain speckle noise due to the data gathering method. Speckle noise is a negative factor for ultrasound images because it complicates the diagnosis of diseases by doctors. Many methods have been developed to overcome this problem [3].

For example, in their study, Gautam and Bharti introduced the bilateral filter method for the enhancement of ultrasound images [1]. Chauhan and Kaushik proposed an anisotropic diffusion filter [4]. Magud et al. used a modified median filter [5]. Gokilavani et al. compared the performances of Gabor and homomorphic fil-

ters [6]. In that study, the Gabor filter was shown to give better results than the other. The wavelet filtering method was introduced by Shanthi and Renuga [7]. An algorithm hybridizing the bilateral filter with NeighShrink was used for improving ultrasound images by Vanithamani and Umamaheswari [8]. Sahu et al. compared the Shock filter, Contrast Limited Adaptive Histogram Equalization, and Spatial filter based on performance metrics [9]. Hafizah and Supriyanto introduced spatial domain filtering, frequency domain filtering, histogram processing, morphological filtering, and wavelet filtering methods. These filters were evaluated based on performance metrics [10].

In contrast to conventional methods, studies are also being performed on ultrasound images with deep learning methods. For example, in the study conducted by Perdios et al., a stacked denoising autoencoder network was introduced to reduce the noise of ultrasound images, and it was shown that the proposed network was superior to the compressed sensing method [11]. In the study conducted by Kokil and Sudharson [12], a pre-trained residual learning network was used for despeckling and it was concluded that this network was superior to classical despeckling filters. In that study, the peak signal-to-noise ratio (PSNR) and structural

E-mail addresses: [onurkaraoglu88@gmail.com](mailto:onurkaraoglu88@gmail.com) (O. Karaoglu), [bilge@gazi.edu.tr](mailto:bilge@gazi.edu.tr) (H. Sakir Bilge), [ihsanuluer@gmail.com](mailto:ihsanuluer@gmail.com) (I. Uluer)



**Fig. 1.** Dilated Convolution Autoencoder Denoising Network schema.

similarity index (SSIM) values of a noisy image and the recommended method's image were 17.71 dB and 0.45 versus 24.10 dB and 0.68, respectively, and the distribution of the noisy image was Gaussian.

Unlike other studies, a generative adversarial network was introduced by Mishra et al. [13]. That study aimed to reduce the speckle noise without disrupting the image's structural features. The proposed deep learning network was very successful compared to other filters.

Speckle noises with Rayleigh distribution at sigma = 0.10, 0.25, 0.50, 0.75 standard deviations were added to ultrasound images obtained from the body's brachial plexus region. The speckle noise model follows Rayleigh distribution [14]. This study aimed to reduce the ultrasound images' speckle noise by using five different deep learning networks (Dilated Convolution Autoencoder Denoising Network/Di-Conv-AE-Net, Denoising U-Shaped Net/D-U-Net, BatchRenormalization U-Net/Br-U-Net, Generative Adversarial Denoising Network/DGan-Net, and CNN Residual Network/DeR-Net). The parameters for block-matching and 3D filtering (BM3D) were obtained from a previous work [15].

In the present paper we propose five deep learning networks that utilize different network structures to reduce the speckle noise content of ultrasound images. The networks proposed here are as follows: a CNN-based autocoder network that consists of dilated convolution layers, U-shaped networks modified by batch normalization and batch re-normalization layers, a residual connection-based denoising network, and a modified generative adversarial network for denoising. The speckle of the ultrasound image dataset is multiplicative noise and the distribution of speckle is Rayleigh. In particular, the proposed networks can handle a wide range of noise levels. For investigating the efficiency of network structures and hyperparameters for denoising problems, the denoising performances of the proposed deep learning networks are compared to each other and to the results of other studies in the literature with similar network structures. Our models are found to outperform other methods, including similar network structures. Our contributions can be summarized as follows:

- We propose five different network structures that can handle a wide range of levels of speckle noise with Rayleigh distribution.
- The denoising performances of the proposed deep learning networks are compared to each other and to other studies in the literature with similar network structures to demonstrate the impact of hyperparameters and network structure.
- Our methods can be applied to real-world noisy images and also show outstanding performance compared to other methods.

In the second part of this paper, the deep learning network structures are explained. Rayleigh distributed noise models, per-

formance evaluation metrics, and test results are presented in Section 3. A discussion and conclusion are provided in Sections 4 and 5, respectively.

## 2. Methodology

In this section, information about the proposed denoising deep learning networks is given. These network structures are chosen from among the most popular deep learning architectures.

### 2.1. Dilated convolution autoencoder denoising network (Di-Conv-AE-Net)

$$F : Z^2 \rightarrow R, \Omega_r = [-r, r]^2 \cap Z^2, k : \Omega_r \rightarrow R$$

If filter size is  $(2r + 1)^2$ , convolution is defined by the following equation [16]:

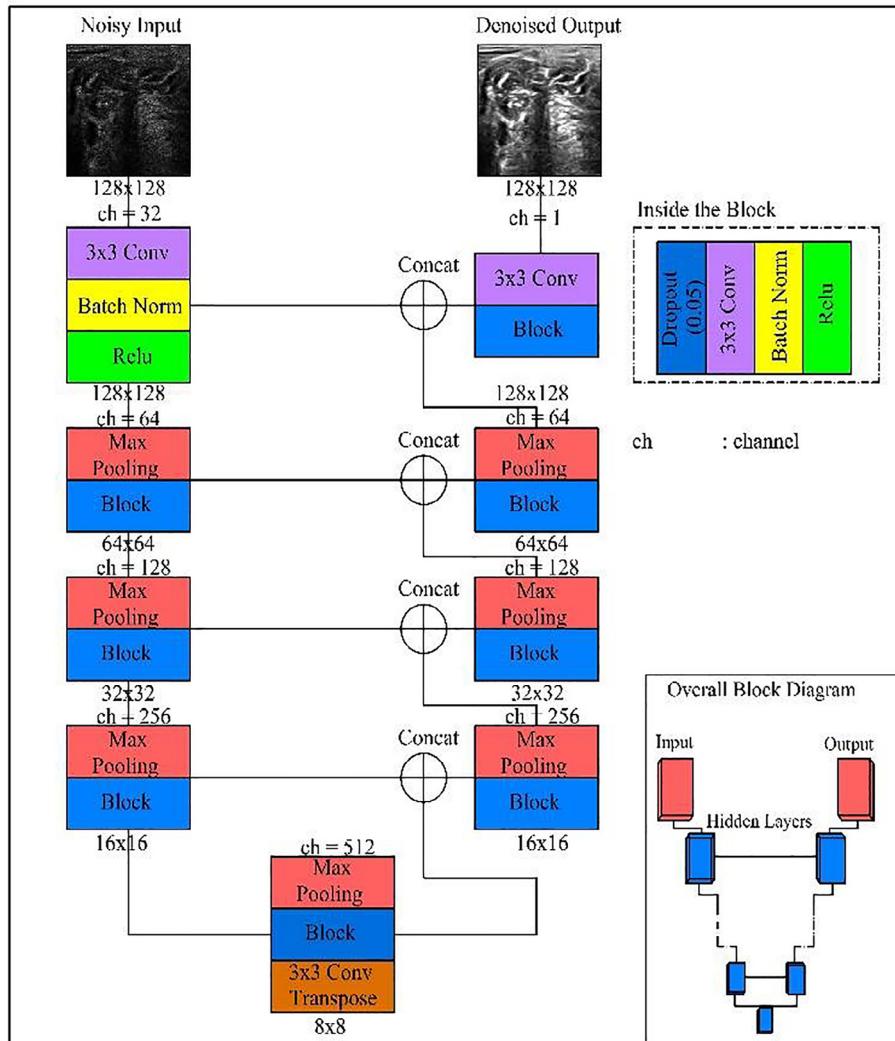
$$(F * k)(p) = \sum_{s+t=p} F(s)k(t) \quad (1)$$

We can expand Equation (1) and introduce expansion factor  $\ell$  as follows:

$$(F * _\ell k)(p) = \sum_{s+\ell t=p} F(s)k(t) \quad (2)$$

The obtained equation is called dilated convolution. Dilated convolution aims to expand the size of an image to prevent data loss. The dilated convolution process is used to avoid max-pooling/downsampling. In CNN-based networks, processes are performed using subsampling layers for the content of the image. While downsampling layers allow the receptive field to expand, they neglect the image resolution. Pooling layers, as a type of downsampling layer, also cause a decrease in the resolution of the images, although they are used for maintaining invariance, controlling overfitting, and performing feature extraction. However, by downsampling layers, the resolution of features through the layers of a network is gradually lost. The resulting coarse features may overlook the details of small objects that are difficult to recover even with efforts involving skip connections or other techniques. Upsampling layers, which are used for restoring the images, are not successful in fixing the resolution [17].

However, resolution is a very important factor for image denoising applications. In dilated convolution, the receptive field is enlarged with the use of holes. Thus, it is ensured that the image is denoised without losing resolution. Moreover, in dilated convolution, while the filter size increases the number of filter parameters, the number of operations per position remains constant [18].



**Fig. 2.** Denoising U-Shaped Net Network Schema.

Consequently, dilated convolution is an alternative way to use a hole filter to replace the pooling and convolution layer [19]. A network that consists of dilated convolution layers and another network consisting of pooling and convolution layers were compared with each other in [18,19]. It is understood from these studies that the resolution of output images obtained from the network with dilated convolution is better.

The dilated convolution network shown in Fig. 1 consists of two parts: an encoder and decoder. The encoder layer contains three dilated convolution layers at levels 1, 2, and 3, respectively, while the decoder layer has three dilated convolution layers at levels 3, 2, and 1, respectively. In this network structure, the maximum pooling layer is used only once.

## 2.2. Denoising U-Shaped Net (D-U-Net)

In the literature, a U-shaped network model was proposed by Ronneberger et al. [20]. This network has two parts: contraction and expansion (Fig. 2). On the network's left side, called the contraction part, noisy image size is reduced by max-pooling and convolution layers. The reduced image is converted to the original image in the expansion part, the network's right side. The expansion and contraction parts of the network make short-circuit con-

nections with each other. In our study, a U-shaped 49-layer deep learning network is used.

Contrary to the work presented in [20], in addition to the use of batch normalization and dropout layers in our D-U-Net network, transpose convolution layers are also used instead of the upsampling layer in the expansion zone of the network. As is known, the convolution layer is used for feature extraction operations in deep learning networks. As a result of that process, the image size is narrowed according to the convolution window. The transposed convolution process is the reverse of the convolution process and the structural properties of the shrinking image size in this layer are restored with minimal distortion. On the other hand, the upsampling layer restores the dimensions of the shrinking image by pairing it. However, structural distortion occurs in the image while it is being restored.

## 2.3. BatchRenormalization U-Net (Br-U-Net)

This network architecture is the same as that of the U-shaped network. The main difference between Br-U-Net and D-U-Net is that the batch normalization layers in this network have been replaced by batch renormalization layers (Fig. 3). The other difference between them is the leaky ReLU layer used in Br-U-Net. The

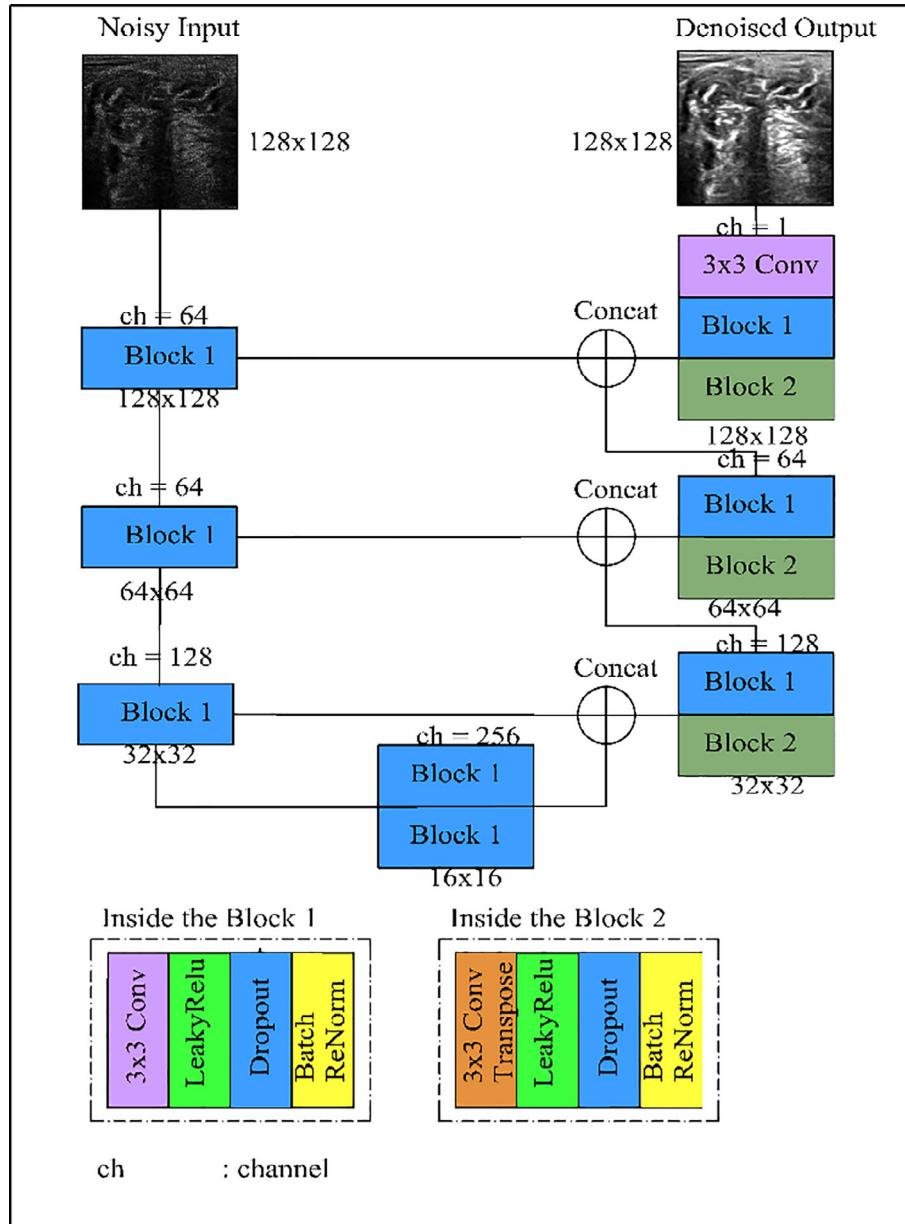


Fig. 3. BatchRenormalization U-Net Network Schema.

batch renormalization layer was proposed by Tian et al. for performing scroll and scale operations to solve the internal variable shifting problem [21]. In the batch renormalization process, the gamma parameter ( $\gamma$ ) and beta parameter ( $\beta$ ) are fixed to 1 and 0.

#### 2.4. Generative adversarial denoising network (DGAN-Net)

In this study, a generative adversarial denoising deep learning network is proposed. This deep learning network was first introduced in 2014 by Goodfellow et al. [22]. It consists of two parts: a generator and discriminator. It is considered that the generator is a counterfeiter. In the denoising GAN network, a noisy image is given as generator input. The denoised image is generated by the generator while the discriminator distinguishes whether the image is real or fake (Fig. 4) [23]. The generator network model is U-shaped and contains 15 layers. The discriminator network also contains 15 layers, and it performs the discriminating process with the help of the convolution layers.

#### 2.5. CNN residual network (DeRNet)

A residual deep learning network is proposed to prevent the vanishing gradient problem [24]. By reducing gradients, shortcut connections are created between layers. Thus, the problem of vanishing gradients is prevented, and the network continues to learn. In our proposed CNN residual network, the layer's size is the same and the number of layers is 25 (Fig. 5).

### 3. Experimental results

#### 3.1. Experimental platform

All experiments are conducted in the Google Colaboratory browser [25]. On this platform, free Nvidia Tesla K80 GPU is offered to users by Google. The Python programming language is used while designing the deep learning network structures, and the Keras and TensorFlow libraries are also used.

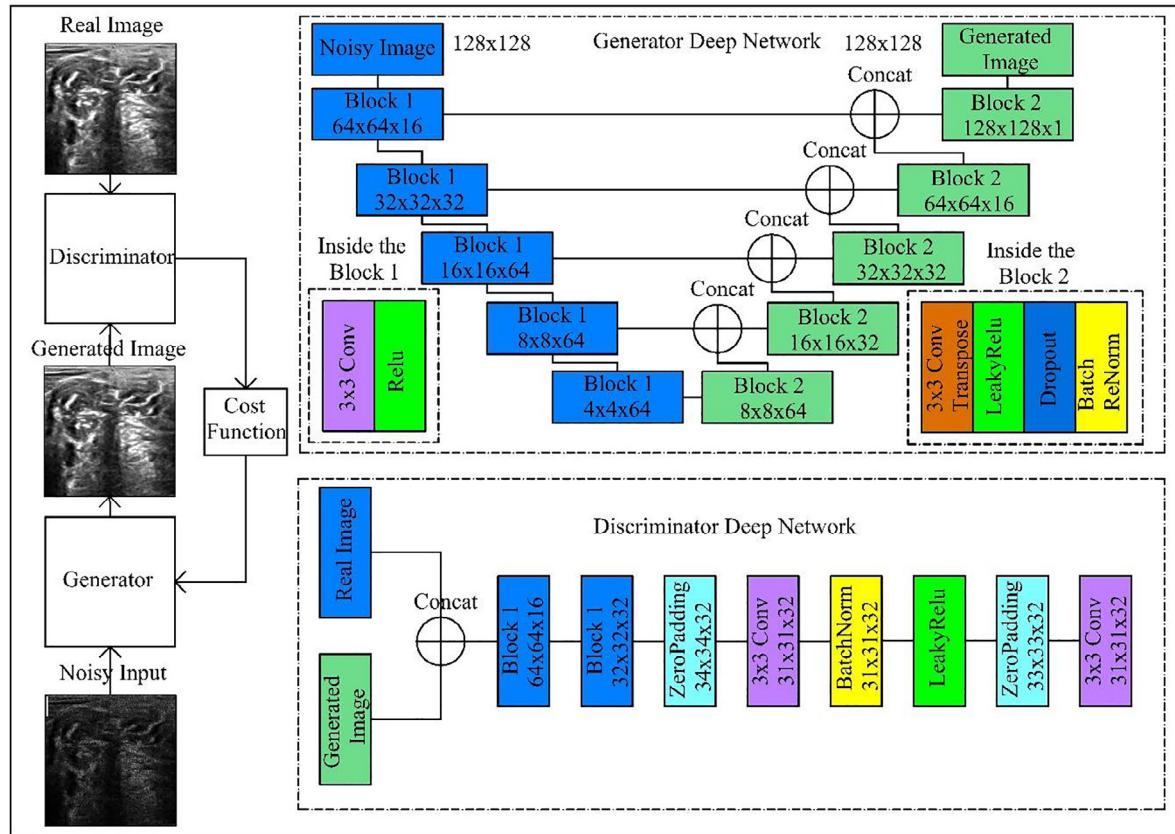


Fig. 4. Generative Adversarial Denoising Network schema.

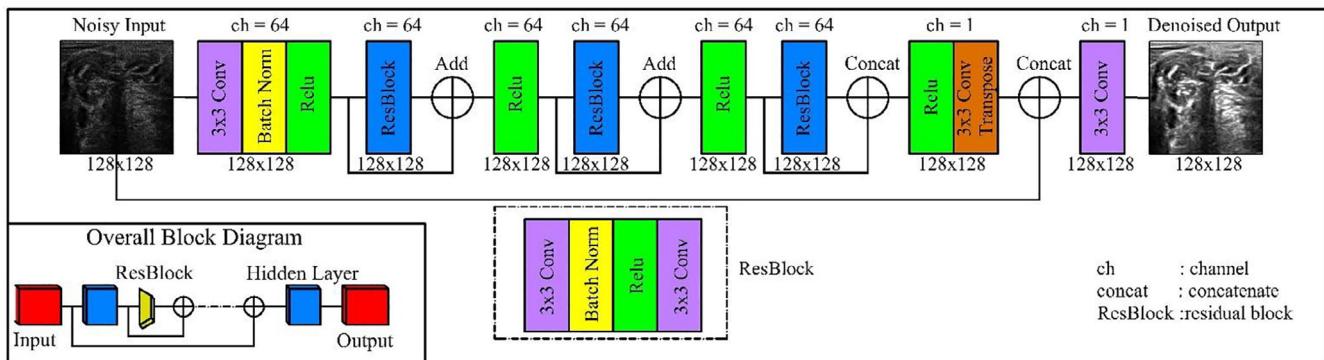


Fig. 5. CNN-Residual Network schema.

### 3.2. Dataset

For the experimental study, the datasets used in the “Ultrasound Nerve Segmentation” competition on the Kaggle site [26] and “Automated Measurement of Fetal Head Circumference” competition [27] are chosen. The number of noisy training images in the first dataset is 2000, the number of clear comparison images is 2000, and the number of test images at 4 different  $\sigma$  levels is 100. Unlike other studies, the training dataset images are adjusted in this study to be equal at all noise levels, rather than adjusted at a single noise level. The first 500 of the 2000 noisy ultrasound images have  $\sigma = 0.10$  standard deviation, the second 500 of the 2000 noisy ultrasound images have  $\sigma = 0.25$  standard deviation, the third 500 of the 2000 noisy ultrasound images have  $\sigma = 0.50$  standard deviation, and the fourth 500 of the 2000 noisy ultra-

sound images have  $\sigma = 0.75$  standard deviation (Fig. 6). The pixel size of the dataset is  $128 \times 128$ .

### 3.3. Adding speckle noise to the dataset

A general mathematical model of speckle noise is expressed as follows [28]:

$$f(x,y) = g(x,y)\eta_m(x,y) + \eta_a(x,y) \quad (3)$$

In Equation (3), the clean image is expressed as  $g(x,y)$ ;  $f(x,y)$  denotes the noisy image and the multiplicative and additive noises are  $\eta_m(x,y)$  and  $\eta_a(x,y)$ , respectively.  $x$  and  $y$  are the spatial positions of the variables. Generally, additive noise is neglected. Radio frequency and envelope signals of backscattered echoes can be



**Fig. 6.** Sample test images from the first dataset.

analyzed using standard stochastic signal theory [29]. Speckle patterns developed for completely random scattering are characterized by Rayleigh distribution.

Assuming that the expression  $\eta_m(x,y)$  in Equation (3) has Rayleigh distribution, the probability density function of  $\eta_m(x,y)$  is expressed as follows:

$$p(f|u) = \frac{f}{u\sigma^2} \exp\left(-\frac{f^2}{2u^2\sigma^2}\right) (\eta \geq 0) \quad (4)$$

The parameter  $\sigma$  represents the noise intensity,  $u$  represents the noise-free image, and  $f$  is the noise-added image.

### 3.4. Performance evaluation

Two quantitative metrics are used to evaluate the performance of the experimental studies. The first one is the peak signal-to-noise ratio (PSNR), which considers the root mean square error (RMSE) between clean and denoised images. The second is the structural similarity index measure (SSIM), which measures the similarity between clean and denoised images.

The equations for the PSNR and SSIM metrics are as follows [30,31]:

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}} \quad (5)$$

Here, MSE represents the mean square error between the real image and predicted image.

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

Here,  $\mu_x$ ,  $\mu_y$ , and  $\sigma_x$ ,  $\sigma_y$  are the mean and standard deviations of the input and denoised images, respectively. The cross-covariance of the image is represented as  $\sigma_{xy}$ .

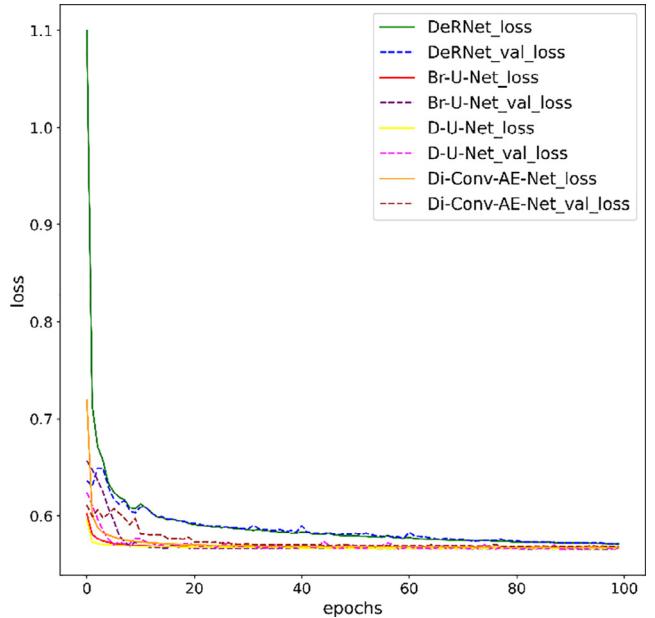
If the PSNR value is high and the SSIM value is close to 1, the applied method's performance is good.

The curves graphed in Fig. 7 demonstrate the loss functions of the four proposed deep learning networks. In this graph, validation loss functions are represented by dashed lines. The generator and discriminator loss functions of DGAN-Net are also presented in Fig. 8. As can be seen from Fig. 7, all loss functions except discriminator loss decrease and balance out after a small number of epochs.

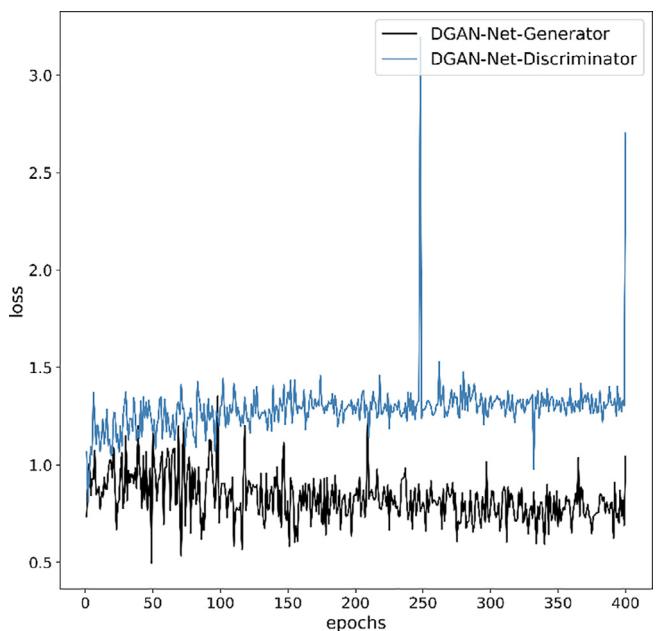
Fig. 8 shows the error function of the generator and discriminator networks in DGAN-Net.

The curves graphed in Fig. 9 demonstrate the accuracy of the four proposed deep learning networks. In this graph, validation accuracies are represented by dashed lines.

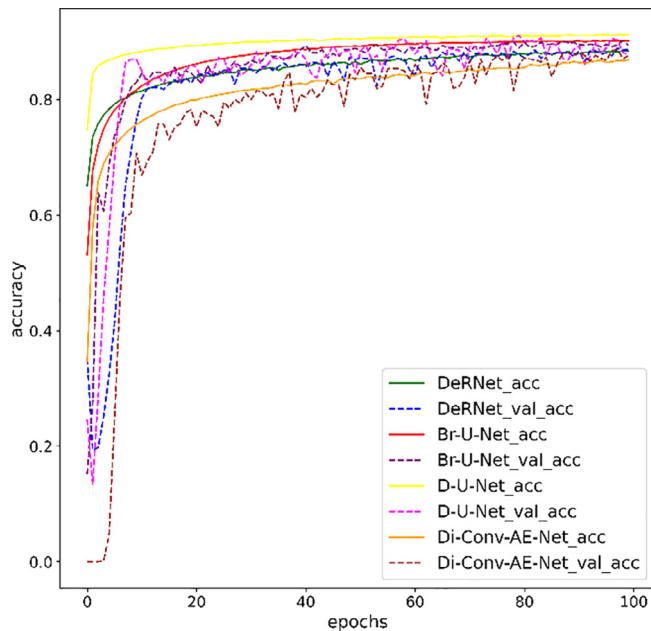
Fig. 10 shows the accuracy of the generator and discriminator networks in DGAN-Net.



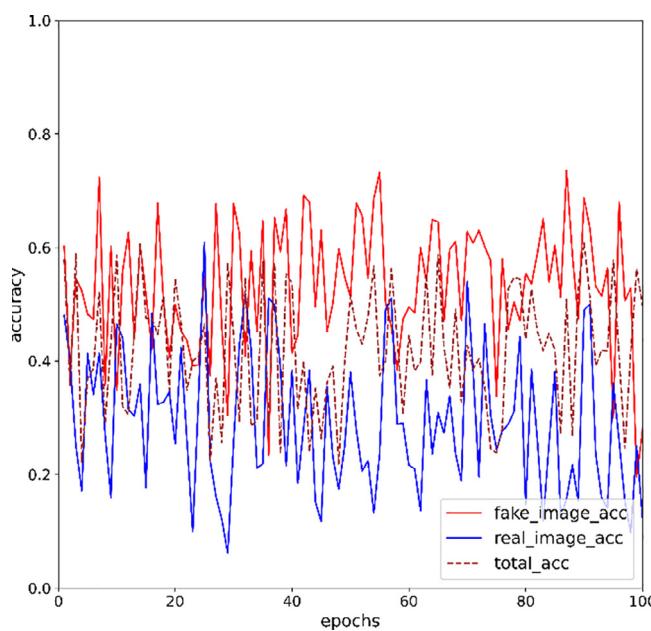
**Fig. 7.** Loss function graphics of deep learning networks, excluding DGAN-Net.



**Fig. 8.** Loss function graphics of DGAN-Net.



**Fig. 9.** Accuracy graphics of deep learning networks excluding DGAN-Net.



**Fig. 10.** Accuracy graphics of DGAN-Net.

### 3.5. Comparisons of five deep learning networks and other denoising methods

In this section, the proposed five deep learning networks are compared with BM3D, classical filtering methods (Bilateral [32], Frost [33], Kuan [34], Lee [35], Mean, and Median Filters), and other deep learning networks (Learning Pixel-Distribution Prior with Wider Convolution for Image Denoising (WIN5-RB) [36], Denoising Prior Driven Deep Neural Network for Image Restoration (DPDNN) [37], and Fingerprint Image Denoising and Inpainting Using M-Net Based Convolutional Neural Networks (PPD-M-Net)) [38]. The results are compared to each other based on the PSNR and SSIM metrics. The hyperparameter values of the selected deep learning networks are also shown. The effects of the hyperparameters on the network results are examined. The classical filter parameters for comparison are determined using the recommended parameter values given in [38].

In our study, four different standard deviations ( $\sigma = 0.10, 0.25, 0.50, 0.75$ ) of speckle noise levels are applied to demonstrate the superiority of the modified deep learning methods compared to other methods for denoising images. The hyperparameters used for the proposed deep learning networks are listed in [Table 1](#). Hyperparameters affect the performance of the deep learning networks and the designer tunes them. If the hyperparameters are changed, the network performance also changes.

In our study, hyperparameters are compared because they can show significant differences between deep learning networks.

Average results of the deep learning networks and the other methods according to PSNR and SSIM metrics are presented in [Table 2](#). Comparisons are made for standard deviation values of 0.1, 0.25, 0.50, and 0.75. It is possible to take standard deviation values as criteria for wider intervals or narrower intervals.

In this study, these standard deviation levels are utilized because their PSNR and SSIM values noticeably differ. As seen in [Table 2](#), PSNR values of the noisy image at standard deviation values of sigma = 0.1, 0.25, 0.50, and 0.75 are 16.62, 13.22, 11.45, and 10.81, respectively, and SSIM values are respectively 87.74%, 69.90%, 53.33%, and 45.90%.

As can be seen from [Table 2](#):

- The network that gives the best PSNR value at 0.10 standard deviation is DGAN-Net with 34.32 dB, and the best SSIM value is 98.69% with U-Net.
- The network that gives the best PSNR value at 0.25 standard deviation is Di-Conv-AE-Net with 32.33 dB, and the best SSIM value is 96.66% with D-U-Net.
- The network that gives the best PSNR value at 0.50 standard deviation is Di-Conv-AE-Net with a value of 31.45 dB, and the best SSIM value is 95.07% with D-U-Net.
- The network that gives the best PSNR value at 0.75 standard deviation is Di-Conv-AE-Net with a value of 29.96 dB, and the best SSIM value is 93.84% with D-U-Net.

**Table 1**  
Deep learning network hyperparameters.

Hyperparameters	Di-Conv-AE-Net	DGAN-NET		D-U-NET	Br-U-NET	DeRNet
		Generator	Discriminator			
Optimizer	Adamax	Adam	Adam	Adamax	Adamax	Adamax
Validation Rate	10%	-	-	10%	10%	10%
Dropout Rate	20%	50%	-	10%	10%	-
Dilated Convolution	Used	-	-	-	-	-
Number of Max-Pooling Layers	1	-	-	4	-	-
Activation Function	Tanh	Tanh	Leaky ReLU	Sigmoid	Sigmoid	Sigmoid
Number of Layers	21	34	11	57	47	24
Learning Rate	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Mini-batch size	20	20	20	32	20	32

**Table 2**

Average first test dataset results of deep learning networks and other denoising algorithms according to PSNR and SSIM criteria.

DENOISING METHODS	PSNR (dB)				SSIM				Time per Epoch (Seconds)
	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75	
Noisy Image	<b>16.62</b>	<b>13.22</b>	<b>11.45</b>	<b>10.81</b>	<b>87.74%</b>	<b>69.90%</b>	<b>53.33%</b>	<b>45.90%</b>	
Bilateral	18.31	15.35	13.92	12.22	73.76%	64.20%	57.96%	47.21%	0.02
Frost	16.9	14.74	13.54	12.02	62.34%	56.40%	52.19%	44.20%	1.40
Kuan	19.1	16.01	14.98	12.67	80.77%	74.96%	70.75%	60.95%	1.65
Lee	19.22	16.07	14.54	12.7	81.65%	75.88%	71.62%	61.77%	0.65
Mean	15.77	13.98	13.34	11.96	53.04%	47.67%	44.00%	38.05%	0.20
Median	16.62	13.22	12.87	11.58	53.04%	47.71%	43.27%	36.93%	0.49
BM3D	21.71	16.79	14.67	13.06	64.07%	63.49%	58.47%	50.66%	31.17
WIN5-RB	32.16	26.78	26.55	25.71	97.08%	94.31%	92.50%	91.17%	21
DPDNN	33	29.26	27.53	26.43	98.36%	96.09%	93.72%	92.20%	127
FDP-M-Net	30.41	28.65	26.58	25.47	97.31%	95.27%	91.46%	89.41%	10
Di-Conv-AE-Net	33.33	<b>32.33</b>	<b>31.45</b>	<b>29.96</b>	96.60%	95.63%	94.26%	93.13%	8
DGAN-Net	<b>34.32</b>	28.20	29.27	27.30	97.87%	95.07%	92.66%	90.67%	3.93
D-U-NET	33.02	27.67	27.45	26.7	<b>98.69%</b>	<b>96.66%</b>	<b>95.07%</b>	<b>93.84%</b>	9
Br-U-NET	31.88	26.5	26.83	25.88	98.54%	96.51%	94.93%	93.54%	15
DeRNet	31.42	27.59	27.3	27.36	96.76%	94.95%	94.16%	93.40%	13

The other denoising methods perform less successfully than the modified deep learning algorithms. As the standard deviation levels increase, the dilated convolution autoencoder has better performance than the other proposed networks and denoising methods. While the Lee filter gives the best PSNR results at 0.10, 0.25, and 0.75 noise levels among the classical filters, the Kuan filter gives the best PSNR result at 0.50. It is indisputable that the Lee filter gives the best results according to SSIM at all sigma levels. However, the Lee filter shows lower performance than BM3D, the other deep learning networks, and the modified deep learning methods trained in this study. Although the BM3D algorithm is one of the most preferred classical filtering methods, its performance is poor compared to the performances of deep learning networks.

Deep learning networks from the literature (DPDNN, WIN5-RB, and FDP-M-Net) are also used to compare the results. These networks used for comparing results have different deep learning network architectures.

The WIN5-RB deep learning architecture is a kind of modified residual network architecture and it has one skip connection between the first and last layers. Therefore, this network and the DeRNet network structure, which is most similar to the WIN5-RB architecture, are compared here in terms of PSNR and SSIM. It is seen that even if the WIN5-RB network is superior to the DeRNet network at the 0.10 sigma level, with increasing sigma levels of noise, the DeRNet network gives better performance.

The DPDNN deep learning architecture is a kind of modified convolution neural network autoencoder architecture. Therefore, this network and the Di-Conv-AE-Net network structure, which is most similar to the DPDNN architecture, are compared here in terms of PSNR and SSIM. It is observed that Di-Conv-AE-Net gives the best results at all sigma levels according to the PSNR metric. According to the SSIM metric, Di-Conv-AE-Net gives the best results except for 0.1 and 0.25 sigma levels.

The FDP-M-Net deep learning architecture is a kind of modified U-shaped deep learning architecture. Therefore, this network and the D-U-Net and Br-U-Net network structures, which are most similar to the FDP-M-Net architecture, are compared according to PSNR and SSIM metric results. As observed from Table 2, the U-Net and Br-U-Net networks give better PSNR results, except at sigma = 0.50, compared to this network. According to the SSIM metric, the U-Net and Br-U-Net networks outperform FDP-M-Net at all sigma levels.

Run times of the deep learning networks and the BM3D algorithm are also compared in Table 2. The DGAN-Net network has the fastest run time compared to the other networks.

Fig. 8 shows the noise-reduced ultrasound images obtained with the deep learning networks and other denoising methods at a standard deviation of sigma = 0.75. In this figure, both noisy images and original images are provided to better illustrate the differences between them. As shown in Fig. 8, the deep learning methods reduce the speckle noise in the images and provide denoised images closest to the real images without disrupting the structural form of the images. Fig. 8 shows how effective these image enhancement methods are by providing the original image and the speckle noise. It is seen that the images obtained with the BM3D algorithm and classical filters are structurally less similar to the original images. This causes difficulty for doctors in diagnosing diseases.

As seen in Table 3, PSNR values of the noisy image at standard deviation values of sigma = 0.1, 0.25, 0.50, and 0.75 are 23.52, 19.79, 17.82, and 16.94, respectively, and SSIM values are respectively 91.84%, 79.75%, 67.95%, and 61.02%.

As can be seen from Table 3:

- The network that gives the best PSNR value at 0.10 standard deviation is DGAN-Net with 35.88 dB, and the best SSIM value is 97.17% with DeRNet.
- The network that gives the best PSNR value at 0.25 standard deviation is DGAN-Net with 33.90 dB, and the best SSIM value is 94.36% with Br-U-Net..
- The network that gives the best PSNR value at 0.50 standard deviation is DGAN-Net with a value of 30.79 dB, and the best SSIM value is 91.64% with Br-U-Net.
- The network that gives the best PSNR value at 0.75 standard deviation is Di-Conv-AE-Net with a value of 29.39 dB, and the best SSIM value is 88.07% with Br-U-Net.

## 4. Discussion

### 4.1. Comparison of proposed networks to each other

As can be seen from the standard deviations, when an image's noise ratio is high, the PSNR value is highest for the dilated convolutional autoencoder network (Di-Conv-AE-Net). The main reason for this is that, unlike other networks, dilated convolution is used here.

Better performance is achieved by providing a minimal reduction in image size by using dilated convolution instead of reducing the image's size. Thus, the best result is achieved without sacrificing image quality.

**Table 3**

Average second test dataset results of deep learning networks and other denoising algorithms according to PSNR and SSIM criteria.

DENOISING METHODS	PSNR (dB)				SSIM				Time per Epoch (Seconds)
	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75	
Noisy Image	<b>23.52</b>	<b>19.79</b>	<b>17.82</b>	<b>16.94</b>	<b>91.84%</b>	<b>79.75%</b>	<b>67.95%</b>	<b>61.02%</b>	
Bilateral	25.12	20.51	20.04	19.41	84.81%	73.96%	72.39%	69.49%	0.03
Frost	23.47	20.02	19.63	19.06	83.85%	73.07%	71.32%	68.34%	1.55
Kuan	27.51	21.55	20.98	20.18	92.10%	81.94%	80.27%	76.35%	1.68
Lee	27.77	21.63	21.03	20.20	92.43%	82.29%	80.46%	76.41%	0.70
Mean	22.5	19.27	18.79	18.22	76.79%	69.34%	67.98%	65.65%	0.25
Median	22.19	19.27	18.79	18.22	77.59%	67.44%	64.80%	61.38%	0.75
BM3D	21.71	16.79	14.67	13.06	64.07%	63.49%	58.47%	50.66%	39.80
WIN5-RB	29.42	27.17	25.4	23.37	96.74%	93.67%	90.44%	87.22%	14
DPDNN	30.22	26.14	23.65	22.06	96.75%	92.25%	86.99%	82.93%	45
FDP-M-Net	26.46	25.8	24.73	22.72	88.99%	85.86%	82.70%	79.73%	4
Di-Conv-AE-Net	33.77	32.22	30.78	<b>29.39</b>	93.97%	91.88%	88.98%	86.04%	5
DGAN-Net	<b>35.88</b>	<b>33.90</b>	<b>30.79</b>	28.92	95.79%	93.29%	89.32%	85.90%	3.50
D-U-NET	28.02	28.66	27.97	25.44	93.80%	92.50%	90.22%	87.27%	3
Br-U-NET	29.42	28.21	27.53	24.98	96.12%	<b>94.36%</b>	<b>91.64%</b>	<b>88.07%</b>	6
DeRNet	31.08	27.45	24.62	22.49	<b>97.17%</b>	93.78%	89.25%	84.68%	10.4

The DGAN-Net deep learning network gives the best PSNR result at the 0.10 level, where the noise is minimal. A generative adversarial network is a deep learning network that can predict an image without seeing the original one. In this study, it was initially assumed that DGAN-Net would perform better at standard deviation values of 0.1 and lower because the PSNR values of other networks decrease as the noise level decreases.

For the SSIM metric, it is seen that the U-Net network gives the best results at all standard deviation values of noisy images. When the D-U-Net network is considered from this perspective, it gives results most similar to the real images.

For example, while the PSNR value of the dilated convolution network at 0.75 sigma noise level is 29.96, the noise level at 0.10 sigma noise level is 33.33. As the noise level decreases, the amount of increase in the PSNR value of this network decreases. However, the PSNR value of DGAN-Net increases at the 0.10 level and reaches 34.32 dB.

According to Table 1, increasing the number of layers of the network does not always positively affect the success of the network. The reason for this is that an increased number of layers causes a vanishing gradient problem.

In addition to visual quality, another important aspect for image denoising methods is the runtime of the networks. Tables 2 and 3 show the runtime performances of the proposed networks. In these tables, runtime is shown as the elapsed time in each epoch of the deep learning network. Although it is seen from Table 2 that DGAN-Net is the fastest deep learning network for each epoch, Di-Conv-AE-Net is the fastest trained network considering the total number of epochs for the network to achieve the desired result. With the longest training time, Br-U-Net is a U-shaped network created without the use of max-pooling layers; therefore, it has the slowest training time compared to other networks as it contains more parameters.

Although it is seen from Table 3 that D-U-Net is the fastest deep learning network for each epoch, DGAN-Net is the fastest trained network considering the total number of epochs for the network to achieve the desired result. With the longest training time, DeRNet has the slowest training time compared to other networks.

#### 4.2. Comparison of proposed networks and other deep learning methods in the literature

In Table 2, the networks recommended in our study are compared to other deep learning network methods that have different network architectures according to PSNR, SSIM, and runtime crite-

ria. Comparisons are made while ensuring that the networks being compared have similar network architectures.

It is understood from Tables 2 and 3 that when WIN5-RB, which has a residual network structure, is compared to the DeRNet network, the DeRNet network performs better in terms of PSNR and SSIM metrics as the noise level increases. Runtime is also an important criterion for comparing deep learning network performances, and although the DeRNet network has more layers than WIN5-RB, it appears to be superior to the WIN5-RB network in terms of training speed. This is because the convolution process is carried out in the WIN5-RB network using a  $7 \times 7$  convolution window.

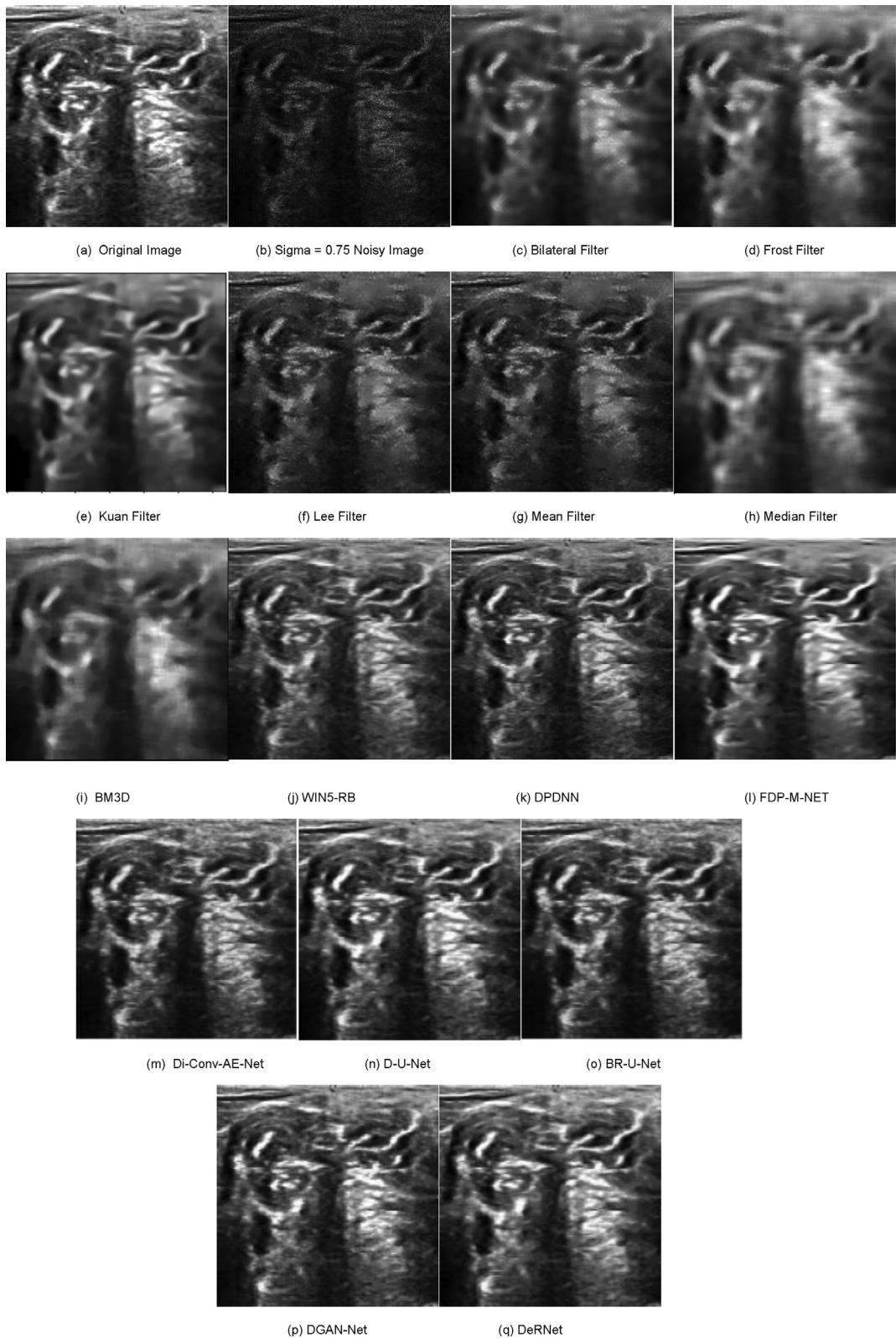
Comparing the DPDNN and Di-Conv-AE-Net networks, which both have convolutional neural network autoencoder architectures, it is understood from Tables 2 and 3 that the Di-Conv-AE-Net network being recommended in our study performs best at all noise levels in terms of PSNR and SSIM metrics. It also outperforms the DPDNN network in terms of its runtime. In addition to having 34 layers, the DPDNN network has a very slow training rate because the number of parameters used in the network is approximately twice that of the Di-Conv-AE-Net network.

It is understood from Tables 2 and 3 that the D-U-Net and Br-U-Net networks recommended in our study perform better at almost all noise levels according to PSNR and SSIM compared to FDP-M-Net, all having U-shaped network structures. The D-U-Net network can be trained faster than FDP-M-Net in terms of runtime. This is because FDP-M-Net consists of 81 layers. The main reason why Br-U-Net has a slower runtime performance than the other networks is that the number of parameters for this network is twice as high as that of the other two networks due to the fact that max-pooling layers are never used in this network.

In summary, the networks recommended in our study outperformed those presented previously in the literature with similar network architectures in terms of PSNR, SSIM, and runtime criteria. The main reasons for this superiority are the creation of networks using sufficiently few layers to minimize vanishing gradient problems, the use of pooling layers at reasonable levels in the proposed networks to minimize image losses, and hyperparameter settings designed to improve the denoising performance of the recommended networks.

#### 4.3. Comparison of proposed networks with classical filters

The networks being recommended in our study are also compared with classical filtering methods and the BM3D algorithm as specified in Table 2 and Table 3. In terms of the PSNR and SSIM



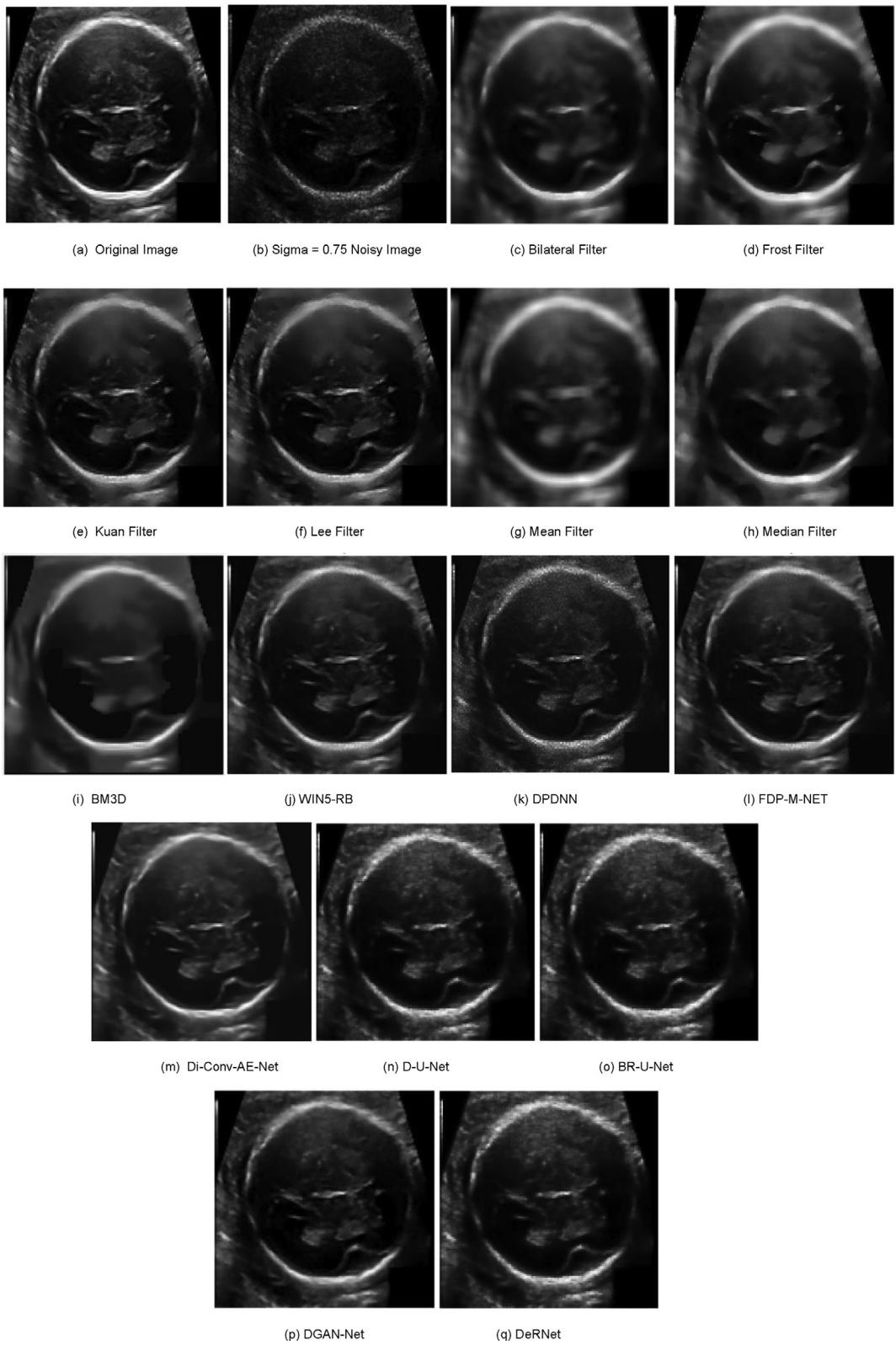
**Fig. 11.** Effectiveness of the proposed deep learning network and the other denoising methods for a standard test image at sigma = 0.75.

metrics, it is clear that the recommended networks show superior performance compared to classical methods.

It is understood from Tables 2 and 3 that the recommended networks provide results faster than BM3D in terms of runtime. Since the number of processes that classic filters apply to reduce noise in images is smaller than the number for the deep learning methods

recommended in our study, our proposed networks require more time for training compared to the classic filters.

As a result, in terms of noise reduction performance, runtime, processes, and number of parameters, the networks that we recommend in this study are more successful than those in the literature and classical filtering methods with similar architectures.



**Fig. 12.** Effectiveness of the proposed deep learning network and the other denoising methods for a standard test image at sigma = 0.75.

We illustrate visual results of Table 2 in Fig. 11 and visual results of Table 3 in Fig. 12. It is clear from Figs. 11 and 12 that the proposed deep learning networks enable the removal of speckle noise from images and increase image resolution. The classical methods and BM3D method are particularly observed to blur the images while denoising the speckle noise in the images.

It is necessary to mention some limitations of this study. First, speckle noise has been added to the datasets because it is impossible to obtain clean ultrasound images under current conditions. Second, deep learning networks perform better when trained with high-resolution images. However, it should not be forgotten that the processing of high-resolution images in deep learning net-

works is only possible with high computational capacity. Third, the proposed networks have been designed only for 2-dimensional B-mode ultrasound images; therefore, there are no data for color or 3D images.

## 5. Conclusions

Ultrasound imaging plays a vital role in the diagnosis of diseases. For doctors to diagnose diseases correctly, speckle noise must be eliminated from ultrasound images. Many filters and methods have been developed to date with that goal. With the increase of computer processor capacities, studies on deep learning, a sub-branch of machine learning, have accelerated. In the present study, deep learning networks, which have more usage areas with the development of computer technologies, are used to reduce speckle noises in ultrasound images.

The speckle noise reduction performances of deep learning networks with five different architectures are compared in this work with BM3D, which is one of the most preferred classical image enhancement algorithms, as well as classical filters including Bilateral, Frost, Kuan, Lee, Mean, and Median Filters and deep learning networks including WIN5-RB, DPDNN, and FPD-M-Net. The performances of the deep learning networks created for this study are superior when compared to the performances of BM3D and classical filters in terms of the PSNR and SSIM metrics. When the deep learning networks are evaluated among themselves, the best PSNR values at 0.25, 0.50, and 0.75 noise levels are achieved with the dilated convolution autoencoder network. The generative adversarial network has the best PSNR performance at the 0.10 noise level. The D-U-Net network has the best SSIM values at all noise levels. This study shows that the proposed deep learning networks are ideal methods for despeckling ultrasound image noises.

With the evaluation of the denoising deep learning networks used in our study and the deep learning networks selected for comparison, it is found that the networks proposed in this study performed better, with the exceptions stated in Section 3.5.

This study also reveals that the creation of the dataset with four different noise levels has a positive effect on network performance. In future studies, the effect of distributing noise levels at different rates rather than utilizing an equal distribution will be examined.

## Declaration of Competing Interest

The 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.

## References

- [1] Gautam, Rekha, and Ms. Rupali Bharti. "Liver Ultrasound Image Enhancement Using Bilateral Filter" International Journal of Engineering and Technical Research (2018),8,4.
- [2] Kamran Binaee, Reza PR Hasanzadeh, An Ultrasound Image Enhancement Method Using Local Gradient Based Fuzzy Similarity, Biomed. Signal Process. Control vol:13 (2014) 89–101.
- [3] S.H.C. Ortiz, T. Chiu, M.D. Fox, Ultrasound Image Enhancement: A Review, Biomed. Signal Process. Control 7 (2012) 419–428.
- [4] Poonam Chauhan\* and Vikas Kaushik, "A Proposed Approach to De-speckle an Ultrasound Image using Anisotropic Filter by Generating Diffusion Tensor", International Journal of Current Engineering and Technology, (2018),8,3
- [5] Ognjen Magud, Eva Tuba, Nebojsa Bacanin, An Algorithm For Medical Ultrasound Image Enhancement By Speckle Noise Reduction, International Journal of Signal Processing 1 (2016) 146–151.
- [6] Gokilavani C., Saravanan M., Kiruthikapreetha.R, Mercy.J, Lawany.R ve Nashreenbanu. M, "Analysis Of Gabor Filter And Homomorphic Filter For Removing Noises In Ultrasound Kidney Images", International Journal of Electrical and Electronics Engineers,2016,8(2), 513-524.
- [7] M. Shanthi, Renuga M."Performance Analysis of Image Enhancement Techniques for kidney Image, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 5 (5) (2016) 3517–3522.
- [8] R. Vanithamani, G. Umamaheswari, Speckle Reduction In Ultrasound Images Using Neighshrink And Bilateral Filtering, Journal of Computer Science 10 (4) (2014) 623–631.
- [9] S. Sahu, M. Dubey, M.I. Khan, Liver Ultrasound Image Analysis using Enhancement Techniques, International Journal of Advanced Computer Research 2 (4) (2012) 125–129.
- [10] W.M. Hafizah, E. Supriyanto, Comparative Evaluation of Ultrasound Kidney Image Enhancement Techniques, International Journal of Computer Applications 21 (7) (2011) 15–19.
- [11] Perdios, Dimitris, et al. "A Deep Learning Approach To Ultrasound Image Recovery". In: 2017 IEEE International Ultrasonics Symposium (IUS). Ieee, 2017. p. 1–4.
- [12] Priyanka Kokil, S. Sudharson, Despeckling Of Clinical Ultrasound Images Using Deep Residual Learning, Comput. Methods Programs Biomed. (2020) 105477.
- [13] Deepak Mishra et al., Ultrasound Image Enhancement Using Structure Oriented Adversarial Network, IEEE Signal Process Lett. 25 (9) (2018) 1349–1353.
- [14] Song, Xiao-Yang, et al. "Speckle Suppression For Medical Ultrasound Images Based On Modeling Speckle With Rayleigh Distribution In Contourlet Domain" 2008 International Conference on Wavelet Analysis and Pattern Recognition. 2008, Vol. 1. IEEE.
- [15] Monagi H. Alkinani, Mahmoud R. El-Sakka, Patch-Based Models And Algorithms For Image Denoising: A Comparative Review Between Patch-Based Images Denoising Methods For Additive Noise Reduction, EURASIP Journal on Image and Video Processing 2017 (1) (2017) 1–27.
- [16] Yu, F., & Koltun, V. "Multi-Scale Context Aggregation By Dilated Convolutions". arXiv preprint. (2015), arXiv:1511.07122.
- [17] R. Hamaguchi, A. Fujita, K. Nemoto, T. Imaizumi, S. Hikosaka, Effective Use Of Dilated Convolutions For Segmenting Small Object Instances In Remote Sensing Imagery March) (2018) 1442–1450.
- [18] L.C. Chen et al., DeepLab: Semantic Image Segmentation With Deep Convolutional Nets, Atrous Convolution, And Fully Connected CRFs, IEEE Transactions On Pattern Analysis And Machine Intelligence 40 (4) (2017) 834–848.
- [19] L. Yuhong et al., "CSRNnet: Dilated Convolutional Neural Networks For Understanding The Highly Congested Scenes.", IEEE Conference On Computer Vision And, Pattern Recogn. (2018) 1091–1100.
- [20] Chunwei Tian, Xu. Yong, Wangmeng Zuo, Image Denoising Using Deep CNN With Batch Renormalization, Neural Networks 121 (2020) 461–473.
- [21] Sergey Ioffe, Batch Renormalization: Towards Reducing Minibatch Dependence In Batch-Normalized Models, Advances In Neural Information Processing Systems (2017) 1945–1953.
- [22] Ian Goodfellow et al., Generative Adversarial Nets, Advances In Neural Information Processing Systems (2014) 2672–2680.
- [23] Zailiang Chen et al., DN-GAN: Denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images, Biomed. Signal Process. Control 55 (2020) 101632.
- [24] Kaiming He et al., "Deep Residual Learning For Image Recognition", IEEE Conference On Computer Vision And, Pattern Recogn. (2016) 770–778.
- [25] <https://colab.research.google.com/>
- [26] <https://www.kaggle.com/c/ultrasound-nerve-segmentation/data>
- [27] Thomas L. A. van den Heuvel, Dagmar de Brujin, Chris L. de Korte and Bram van Ginneken, Automated measurement of fetal head circumference using 2D ultrasound images . Zenodo. <http://doi.org/10.5281/zenodo.1322001>.
- [28] Xiao-Yang Song et al., Speckle Suppression For Medical Ultrasound Images Based On Modeling Speckle With Rayleigh Distribution In Contourlet Domain, International Conference on Wavelet Analysis and Pattern Recognition 1 (2008) IEEE.
- [29] T.A. Tuthill, R.H. Sperry, K.J. Parker, Deviations From Rayleigh Statistics In Ultrasonic Speckle, Ultron. Imaging 10 (2) (1988) 81–89.
- [30] P. Khandelwal, G. Kaur, Comparative Study of Different Image Enhancement Technique, International Journal of Electronics & Communication Technology 7 (2) (2016) 116–121.
- [31] Zhou Wang et al., Image Quality Assessment: From Error Visibility To Structural Similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
- [32] Michael Elad, On The Origin Of The Bilateral Filter And Ways To Improve It, IEEE Trans. Image Process. 11 (10) (2002) 1141–1151.
- [33] Victor S. Frost et al., in: IEEE Transactions on pattern analysis and machine intelligence, 1982, pp. 157–166.
- [34] Darwin T. Kuan et al., in: IEEE transactions on pattern analysis and machine intelligence, 1985, pp. 165–177.
- [35] LEE, Jong-Sen, Speckle Analysis And Smoothing Of Synthetic Aperture Radar Images, Computer Graphics And Image Processing 17 (1) (1981) 24–32.
- [36] Liu, Peng, and Ruogu Fang, "Learning Pixel-Distribution Prior With Wider Convolution For Image Denoising," arXiv preprint arXiv: (2017). 1707.09135
- [37] W. Dong, P. Wang, W. Yin, G. Shi, F. Wu, X. Lu, Denoising Prior Driven Deep Neural Network For Image Restoration, IEEE Transactions On Pattern Analysis And Machine Intelligence 41 (10) (2018) 2305–2318.
- [38] Kiran Dasari, L. Anjaneyulu, Importance Of Speckle Filter Window Size And Its Impact On Speckle Reduction In SAR Images, International Journal of Advances in Microwave Technology 2 (2017) 98–102.