A Multiscale Convolutional Neural Network for Speckle Noise Removal in Optical Coherence Tomography Images

M.NagoorMeeral[1], Dr.S.Shajun Nisha[2] & Dr. M.MohamedSathik[3]

[1]Reg No.19211192282029, Ph.D Research Scholar, PG & Research Department of Computer Science, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli, India, Affiliated to Manonmaniam Sundaranar University,Abishekapatti,Tirunelveli-627012,Tamil Nadu, India.

[2]ResearchSupervisor, Assistant Professor & Head, PG & Research Department of Computer Science, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli, Tamil Nadu, India.

[3]Principal, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli,TamilNadu, India.

Abstract

Optical Coherence Tomography (OCT) is proficient in imaging micrometric resolution of retinal cross sections to identify ophthalmology diseases. However, the OCT images are corrupted by multiplicative speckle noise, generates severe degradation in its quality. The challenge is to remove these inherent noises to avoid deceptive results in fringe areas. Recent advances in discriminative learning can promisingly remove complex noises in SD-OCT images. This paper proposes a methodology for effective denoising model which deploys Convolutional network with Multiscale operation. A huge amount of local features can be generated with applying parallel dissimilar filter sizes. The performance of the proposed approach is assessed on Duke (SD-OCT) dataset. The suggested technique is evaluated against conventional speckle denoising methods on parameters namely PSNR, SSIM, AD, LMSE, NK and NAE. Experimental analysis show that our method can effectively minimize noise and preserve the retinal structures than other traditional methods in terms of both quantity and quality assessment.

Keywords: SD-OCT, Speckle denoising, Convolutional Neural Network, Discriminative learning.

M.Nagoor Meeral, Dr. S.Shajun Nisha & Dr. M.Mohamed Sathik

1.Introduction

OCT is a highly versatile screening technology, has been used in variety of applications particularly to facilitate monitoring and earlier diagnosis of ophthalmology diseases like diabetic retinopathy, Glaucoma, Macular degeneration etc. It is highly recommended by ophthalmologists as it owns the important features. It is a non-invasive procedure, capable of acquiring retinal cross sections at a high speed. The principle of low coherence (white light) interferometer is followed, which reduces the probing range to micrometers. It exploits broad-bandwidth light sources, able to capture 10 micrometer or less spatial resolution [1] [2] [3].

The dual property of interferometer in OCT is higher resolution imaging and speckle formation. The spatial and temporal coherence of the scattered signals returning from the image volume provokes an uncorrelated grainy texture known as speckle. Multiple backscattering of these light sources induces Speckle impression in lieu of propagation depth. As it is multiplicative in nature, it comprises both information carrying signal and quality diminishing noise. The signal and noise can be discriminated based on correlation spot size and frequency of waves, but the spatial distribution of these two differs significantly. Therefore, an efficient technique must be followed to accentuate the constructive interference by suppressing the other factor [4] [5].

Speckle reduction is a difficult task since it is correlated with the microscopic information of the retina. To maintain the trade-off between image quality and noise suppression, two different approaches can be adopted: Instrument-based and Image based denoising methods. The former one involves compounding techniques like using separated light with varying bandwidth (frequency compounding) [6], averaging magnitude of signals (spatial compounding) [7], receiving back scattering at distinct angles (angular compounding) [8] and applying polarization state of light (polarization diversity) [9]. These approaches result in limited resolution, system complexity and high acquisition time.

The latter one deploys various filtering techniques in the earlier stages. These methods result in minimal performances when the noise level is increased. The denoising algorithms are categorized into model centric and discriminative learning techniques. Model centric algorithms include anisotropic diffusion (AD) [10], Non-Local Means filter (NLM) [11], Block matching and 3Dfiltering (BM3D) [25] and Weighted Nuclear Norm Minimization (WNNM) [26] are

developed. Usually, these models utilizes the Non-local Self Similarity (NSS) and image priors and achieve considerable SNR value/promising results. However, the two major drawbacks are time consuming and the parameters need to be manipulated every time.

The application of discriminative learning has a huge impact on image restoration problems. Especially Convolutional neural networks are effective in manipulating image features [12] [13].The advances in regularization and learning strategies drive CNN to achieve better training performance. Hence, it gains huge attention over computer vision problems. This research work aims to build a CNN with moderate deep architecture to denoise OCT images. The hidden layers in the network intend to extract the speckle noises from the degraded image, which implies the output will be denoised OCT image. In addition, the network handles multiple convolutions at three different kernel sizes simultaneously. The performance of this research work is validated using metrics like Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Average Difference (AD), Normalized Cross Correlation (NK), Normalized Absolute Error (NAE) and Laplacian Mean Square Error (LMSE).

The following steps summarize the threefold contributions of the proposed Multiscale CNN model:

1. The proposed Multiscale CNN works on downsampling images in which the network can acquire reasonable depth for accelerated data processing.

2. Parallel Convolution filters are employed in the architecture to acquire the distinct patterns of image. The width of the model is increased in a particular convolution specially to get the optimum feature representations.

3. Following experiments with activation functions, it is clear that Leaky ReLU supports faster training. Moreover it allows small negative gradients when the input is less than zero.

1.1. Motivation and Justification

The presence of the granular speckle structure raises complexity in retinal layer segmentation and disease diagnosis. Hence noise suppression remains indispensable task in OCT image analysis. Nowadays, Deep learning techniques are in the mainstream regardless of several

denoising methods are available in the literature. Motivated by this, an efficient deep denoising architecture named Multiscale Convolutional Neural Network to remove Speckle Noise in Optical Coherence Tomography images is proposed here.

According to the literature study, Downsampling of input images help in improving the efficiency of CNN model. It increases the receptive field and supports minimizing the network depth. It is crucial for resolving the issue of performance deterioration and memory burden. Additionally, to enhance the performance of the training accuracy, Multiscale operation is also exploited to extract features using different filter sizes [14] [15]. From the existing research works, it is found that Multiscale operation can augment the signal to noise ratio by preserving the image quality correspondingly. To justify this, pioneering denoising models are used to compete with the proposed one and it is demonstrated in both quantitative and qualitative evaluation.

1.2. Organization of the Paper

The structure of the paper is prepared in the following way: Section 2 comprises Related Work, Section 3 covers Proposed Denoising Model, Section 4 includes Experimental Results, and Section 5 contains Conclusion.

2. Related Work

Numerous models based on deep Convolutional Neural Network have been presented in the literature to explain its potential of handling Noises in Optical Coherence Tomography images. A DnCNN model was proposed by Zang et al. [16] for Gaussian denoiser. In DnCNN, concept of residual learning adding with batch normalization improves the training accuracy. This DnCNN model is a blind denoising model which performs better for three image distinct tasks, i.e., blind Gaussian denoising, Single Image Super Resolution problem (SISR), and JPEG deblocking. DnCNN model is remodeled by Gour et al. [17] which implement the concept of residual learning. The architecture produces a reconstructed image after 17 layers of convolution. The noise residuals are obtained as the result, which is finally subtracted from the original image to get the denoised output image. The performance of the model is evaluated in Topcon and Duke databases.

A Multiscale Convolutional Neural Network for Speckle Noise Removal in Optical Coherence Tomography Images

Zhang et al. [18] introduced a fast and flexible denoising Convolutional Neural Network (FFDNet) which greatly concerns with spatially variant noises. An adjustable noise level map along with the down sampled images are entailed here as input which is a major contribution to maintain an optimal balance between noise reduction and detail preservation. Anoop et al. [19] developed DenseNet based architecture to learn the noise behavior in OCT images and implements patch-wise methodology for despeckling. Shen et al. [20] introduced a double-path parallel Convolutional Neural Network (DPNet) using dilated convolution and residual learning to improve performance. Li et al. [21] proposed a new model based on structural CNN and feature fusion. The network architecture encompasses three sub networks. It utilizes U-Net and several Convolutional layers for speckle noise suppression.

3. Proposed Denoising Model

3.1. Network Architecture

The proposed network adopts FFDNet. The architecture is demonstrated in Figure 1. The input image y=x + v where x is the original image and v is the noise. The architecture encompasses 11 layer depths which perform a series of Convolutional operations, without pooling operations. The Noisy OCT image with size W x H x C is given as Input where C=1 represents gray scale image. It undergoes maxpooling operation with stride 2 for downsampling the image size to ½(W x H) x C. The first layer (Convolution + Leaky ReLu) are performed in the convolution layer including 64 numbers of 3x3 filters. To construct the network more optimized, Leaky ReLu activation is used which accept small positive slopes when node is inactive. The formula for Leaky ReLu activation function is given in Equation1.

$$f(x) = \begin{cases} x & for\ x > 0 \\ 0.01x & otherwise \end{cases} \tag{1}$$

Then the second layer includes Batch Normalization along with Convolution and Leaky ReLu. Multiscale operations are deployed in the third layer, contains parallel convolution with three different filter sizes 3x3, 5x5 and 7x7. The output of 3 parallel convolutions is then concatenated to extract diverse range of features, which are then passed to seven consecutive convolution layers. Figure2 depicts the block diagram of Multiscale Feature Fusion in Proposed Denoising Model. The operation of Multiscale Feature Fusion can be expressed as Equation 2.

$$Y_3^{out} = Conv(Y_2^{out}, K1) \oplus Conv(Y_2^{out}, K2) \oplus Conv(Y_2^{out}, K3) \qquad (2)$$

Where, Y_3^{out} = output of Multiscale Layer; Y_2^{out} is the previous layer output (Layer 2); K1=3x3; K2=5x5; K3=7x7 are the filters. Conv represents the Convolution + BN + Leaky ReLU. The last layer is the convolution with filter of size 3×3×1 to restructures the output. Finally, the output is up sampled to obtain the image original size W x H x C. Table 1 illustrates the summary of proposed Network model.
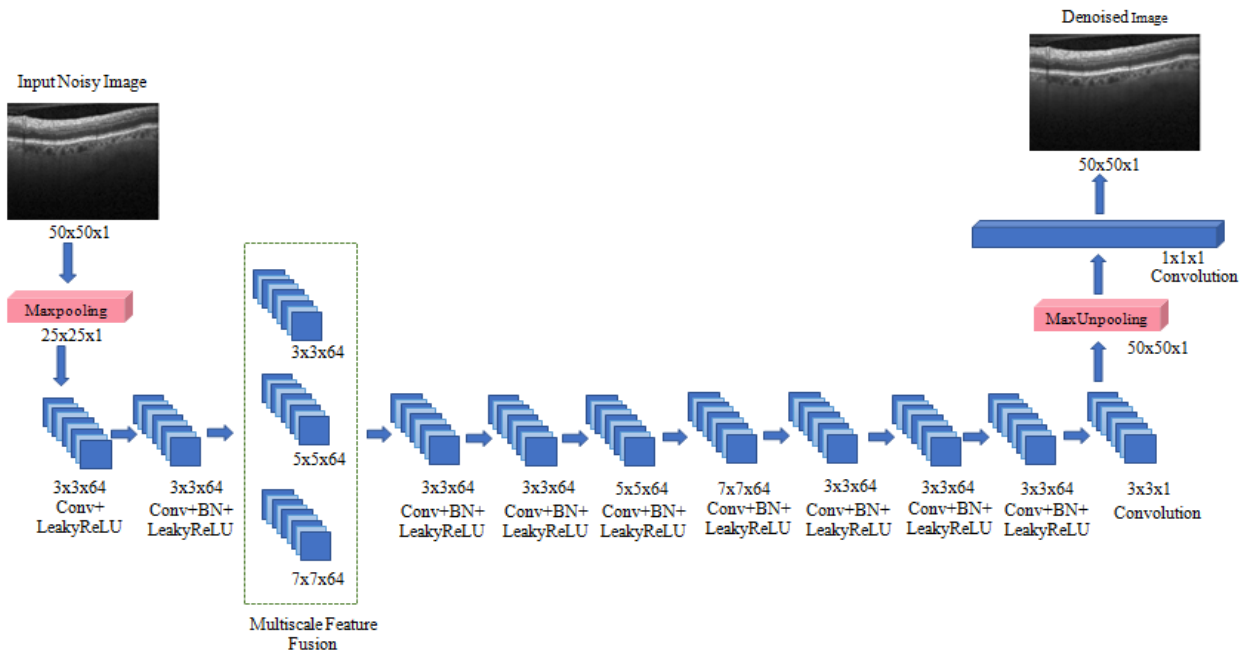
Figure 1. Network Architecture of Proposed Denoising Model
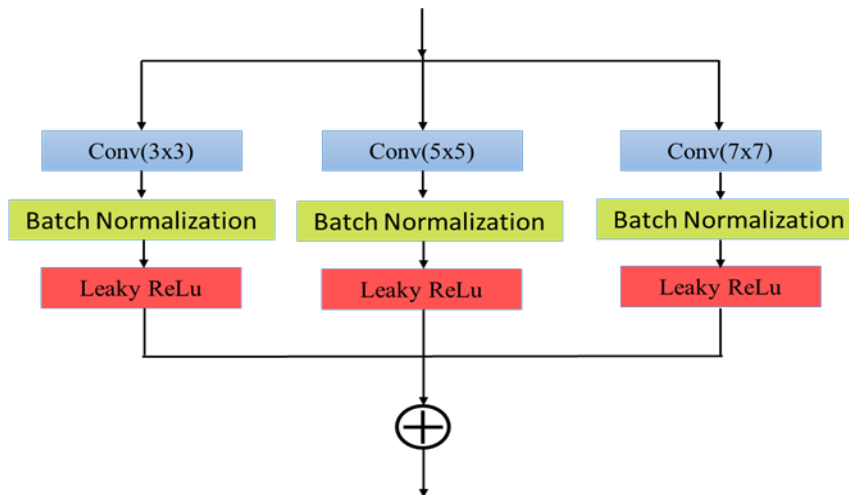
Figure 2. Block diagram of Multiscale Feature Fusion in Proposed Denoising Model

A Multiscale Convolutional Neural Network for Speckle Noise Removal in Optical Coherence Tomography Images

Table 1. Summary of Proposed Network Model

Layer	Module	No. of filters	Kernel size	Stride	Output shape	Parameters
Maxpool	-	-	2x2	2	25x25x1	0
Conv_1	Convolution, Leaky ReLu	64	3x3	1	25x25x64	640
Conv_2	Convolution,BatchNorm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_3	Convolution,Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_4	Convolution,BatchNorm, Leaky ReLu	64	5x5	1	25x25x64	102592
Conv_5	Convolution,Batch Norm, Leaky ReLu	64	7x7	1	25x25x64	200896
Addition	-	-	-	-	25x25x64	0
Conv_6	Convolution,Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_7	Convolution,Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_8	Convolution,Batch Norm, Leaky ReLu	64	5x5	1	25x25x64	102592

Conv_9	Convolution,Batch Norm, Leaky ReLu	64	7x7	1	25x25x64	200896
Conv_10	Convolution, Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_11	Convolution, Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_12	Convolution, Batch Norm, Leaky ReLu	64	3x3	1	25x25x64	37056
Conv_13	Convolution	1	3x3	1	25x25x1	577
MaxUnpool	-	-	-	-	50x50x1	0
Conv_14	Convolution	1	1x1	1	50x50x1	2

Table 2. Network Specifications

	DnCNN	FFDNet	Proposed
Network depth	20	17	11
Activation	ReLu	ReLu	Leaky ReLu
Optimizer	SGD	Adam	Adam
Learning Rate	$10^{-1}$to 10^{-4}	$10^{-3}$to 10^{-4}	10^{-3}
Epoch	50	50	50
Mini-batch Size	128	128	128

4. Experimental Analysis

This section comprises Database description, Network training, Comparison with state-of-the art methods, Performance Metrics, Results and Discussion, Running Time.

4.1. OCT databases

To demonstrate the performance of the competing methods, experiments are carried on Duke dataset. It contains SD-OCT images of 45 patients. The dataset comprises three classes of SD-OCT images namely: 15 Normal patients, 15 Dry AMD and 15 DME. The patient's retinal images are captured from Spectralis SD-OCT diagnostic device (Heidelberg Engineering Inc.) [22]. A total of 600 images are taken from the dataset. Both the pathological and non-pathological images are included for training and testing process.

4.2. Network Training

From the Duke dataset, 480 images are used for training and 120 images are used for testing. Given the input patch size of 50 x 50, five different noise levels are considered, i.e., σ = 5,15,25,50 and 75 to train the proposed network. The training phase is executed for 50 epochs. The Network specifications are described in Table 2.Let, N be the number of noisy-ground truth image pairs (y_i, x_i) for training parameters (θ),then average mean squared error $L(\theta)$calculates the difference between ground truth and the predicted image. It can be defined as,

$$L(\theta) = \frac{1}{2N}\sum_{i=1}^{N}\|F(y_i; \theta) - x_i\|^2 \tag{3}$$

The ADAM algorithm [23] is used as an optimizer to minimize the loss function with momentum of 0.9.The learning rate is10^{-3} and the mini-batch size is set as 128.

4.3. Comparisons with state-of-the art methods

Several state-of-the art methods are compared to prove the efficiency of the proposed Multiscale CNN like Anisotropic diffusion, WNNM, NCSR, BM3D, DnCNN, FFDNet. Anisotropic diffusion is an inhomogeneous technique which reduces noise by preserving the boundaries or edges. It follows non-linear diffusion process [10].

WNNM is an advancement of Nuclear Norm Minimization (NNM). It significantly increases the Signal to Noise Ratio (SNR), without using the prior knowledge.

WNNM allots different weights for different singular values so that the subtle information from the input data will be preserved without any loss [26].Non locally Centralized Sparse Representation (NCSR) was introduced in 2013.It employs non-local self-similarity to reduce sparse coding Noise. Considering iterative shrinkage to reduce minimization problem makes NCSR an effective denoise model [24]. BM3D approach is implemented with two different processes in the 3Dtransform domain: block-matching and collaborative filtering. The Noisy image is processed as blocks using a sliding window. To accomplish collaborative filtering, the identical blocks from the noisy image are retrieved and piled together to create a 3D array. The denoised image is constructed by aggregating several estimations of each pixel [25]. DnCNN is a discriminative learning approach which endeavors residual strategy. DnCNN extracts the residual image using the hidden layers. The combination of Residual learning and batch normalization chiefly responsible for its productive denoising results [16]. FFDNet is CNN based architecture, designed to remove spatially variant noises. It comprises two different strategies like (1) the noise level map is utilized to preserve the image details. (2) The network acts upon downsampled images to possess moderate network depth [18].

4.4. Performance Metrics

For parametric evaluation, Metrics like PSNR, SSIM, AD, NK, NAE, LMSE [27], [28], [29], [30] are used.

Peak Signal to Noise Ratio (PSNR):

PSNR refers Peak to Signal-To-Noise Ratio can be measured in decibels (db). It measures the quality of reconstructed image. The higher the value, the higher its quality. The mathematical formula for PSNR calculation for the original image I and the denoised image K with size m × n is expressed as,

$$PSNR = 10 \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{4}$$

Where, MAX_I is the maximum possible value of the image I.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{5}$$

Structural Similarity Index (SSIM):

SSIM, which is a perception-based model, measures the similarity between the original image and noise free image. Combinations of three features are involved in comparison namely: Luminance, Contrast and Texture. It can be defined as,

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

Where, μ_x, μ_y are the average of x and y. σ_x, σ_y are the variance of x and y and σ_{xy} is the covariance of x and y, c_1 and c_2 are the constants.

Average Difference (AD):

AD calculates the pixel wise difference between the true image and the denoised image. If the value of AD is larger, then the image is of poor quality.

$$AD = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)] \qquad (7)$$

Where, m, n is the image size, I and K are the original and obtained image.

Normalized Cross Correlation (NK):

NK calculates similarity measurement between the compared images. The value falls in the range between -1 and 1.

$$NK = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)]}{\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j)]^2} \qquad (8)$$

Where, I and K are the original and denoised image, m, n is the image size.

Normalized Absolute Error (NAE):

It measures the difference between original and reconstructed image. The result lies between the interval 0 and 1.The greater the value, the lower the similarity. The greater the value, the lower the similarity.

$$NAE = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}|I(i,j) - K(i,j)|}{\sum_{i=1}^{m}\sum_{j=1}^{n}|I(i,j)|} \qquad (9)$$

Where, I and K are the original and obtained image and m, n is the image size.

Laplacian Mean Square Error (LMSE):

LMSE is determined by the Laplacian value of the expected and obtained image. The value of LMSE is inversely proportional to the image quality. The larger the value, the poorer the quality. It is defined as,

$$LMSE = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} [O\{I(i,j)\} - O\{K(i,j)\}]^2}{\sum_{i=1}^{m} \sum_{j=1}^{n} [O\{I(i,j)\}]^2} \tag{10}$$

Where, I and K are the original and denoised image and m, n is the image size.

Table 3. Quantitative Analysis of existing with proposed model for noise level15, 25, 35, 50 and 75

$\sigma=15$						
	PSNR	**SSIM**	**NK**	**NAE**	**LMSE**	**AD**
Anisotropic Diffusion	28.8256	0.7018	0.9895	0.3254	0.9227	7.9518
BM3D	29.925	0.7017	0.9931	0.273	0.8929	5.0428
WNNM	30.5232	0.7197	0.9942	0.1978	0.75	4.1342
NCSR	30.1125	0.73414	0.9946	0.1916	0.68	4.4587
DnCNN	30.5966	0.82	0.9965	0.0666	0.56	4.5988
FFDNet	31.7402	0.8581	0.9974	0.0534	0.4269	3.7419
Proposed Multiscale CNN	**34.3378**	**0.9338**	**0.9986**	**0.0382**	**0.3026**	**2.6912**
$\sigma=25$						
Anisotropic Diffusion	28.5653	0.6978	0.9756	0.2949	0.9701	8.8879
BM3D	29.1858	0.7225	0.9925	0.2785	0.9395	7.8634

WNNM	29.75	0.7432	0.9921	0.2845	0.8867	7.5214
NCSR	29.2431	0.7346	0.99458	0.2678	0.8834	7.46352
DnCNN	30.2044	0.7075	0.9879	0.0997	0.8471	6.797
FFDNet	31.501	0.7615	0.9933	0.0765	0.5733	5.3919
Proposed Multiscale CNN	**33.3927**	**0.8308**	**0.9967**	**0.0596**	**0.4814**	**4.1754**
$\sigma=35$						
Anisotropic Diffusion	26.0245	0.6010	0.9684	0.3546	2.2318	9.9648
BM3D	27.1638	0.6351	0.972	0.2865	1.9435	9.0542
WNNM	28.8912	0.73256	0.97256	0.2416	1.9572	8.1274
NCSR	28.9925	0.73414	0.9933	0.1914	1.7096	8.0502
DnCNN	28.8631	0.7651	0.9917	0.0833	1.0406	6.0406
FFDNet	32.1893	0.8268	0.9959	0.064	0.5643	4.4962
Proposed Multiscale CNN	**33.3756**	**0.8303**	**0.9967**	**0.0597**	**0.4824**	**4.185**
$\sigma=50$						
Anisotropic Diffusion	25.0475	0.5021	0.834	0.4810	2.4281	10.8243
BM3D	25.6153	0.5061	0.8219	0.4562	2.1855	10.7826
WNNM	26.9542	0.5954	0.9517	0.356	1.5437	9.4357
NCSR	26.8415	0.5122	0.9871	0.3153	1.2318	8.6275

DnCNN	28.2608	0.6976	0.9916	0.0957	0.7652	6.5967
FFDNet	28.3085	0.7005	0.9899	0.0823	0.8183	5.9023
Proposed Multiscale CNN	**30.4295**	**0.7807**	**0.9935**	**0.084**	**0.6497**	**5.8883**
$\sigma=75$						
Anisotropic Diffusion	24.0142	0.4452	0.7583	0.6942	2.9854	10.8241
BM3D	24.6814	0.5603	0.7827	0.6137	2.9932	10.533
WNNM	24.359	0.5842	0.894	0.6699	2.6518	10.0538
NCSR	24.1127	0.5813	0.9736	0.5182	2.5417	9.1675
DnCNN	25.6673	0.5023	0.9851	0.1158	1.2423	9.6575
FFDNet	27.5073	0.6818	0.9872	0.1178	0.7865	8.2645
Proposed Multiscale CNN	**27.6393**	**0.7069**	**0.9885**	**0.1141**	**0.7053**	**8.0656**

Table 4.PSNR values of state-of-the-art with proposed model for noise level $\sigma=5$ to $\sigma=50$

Noise Level	DnCNN	FFDNet	Proposed Multiscale CNN
$\sigma=5$	32.5274	32.9313	**34.3392**
$\sigma=10$	32.2742	32.9313	**34.3392**
$\sigma=15$	31.0566	32.702	**34.3338**
$\sigma=20$	30.5988	32.572	**33.9876**

A Multiscale Convolutional Neural Network for Speckle Noise Removal in Optical Coherence Tomography Images

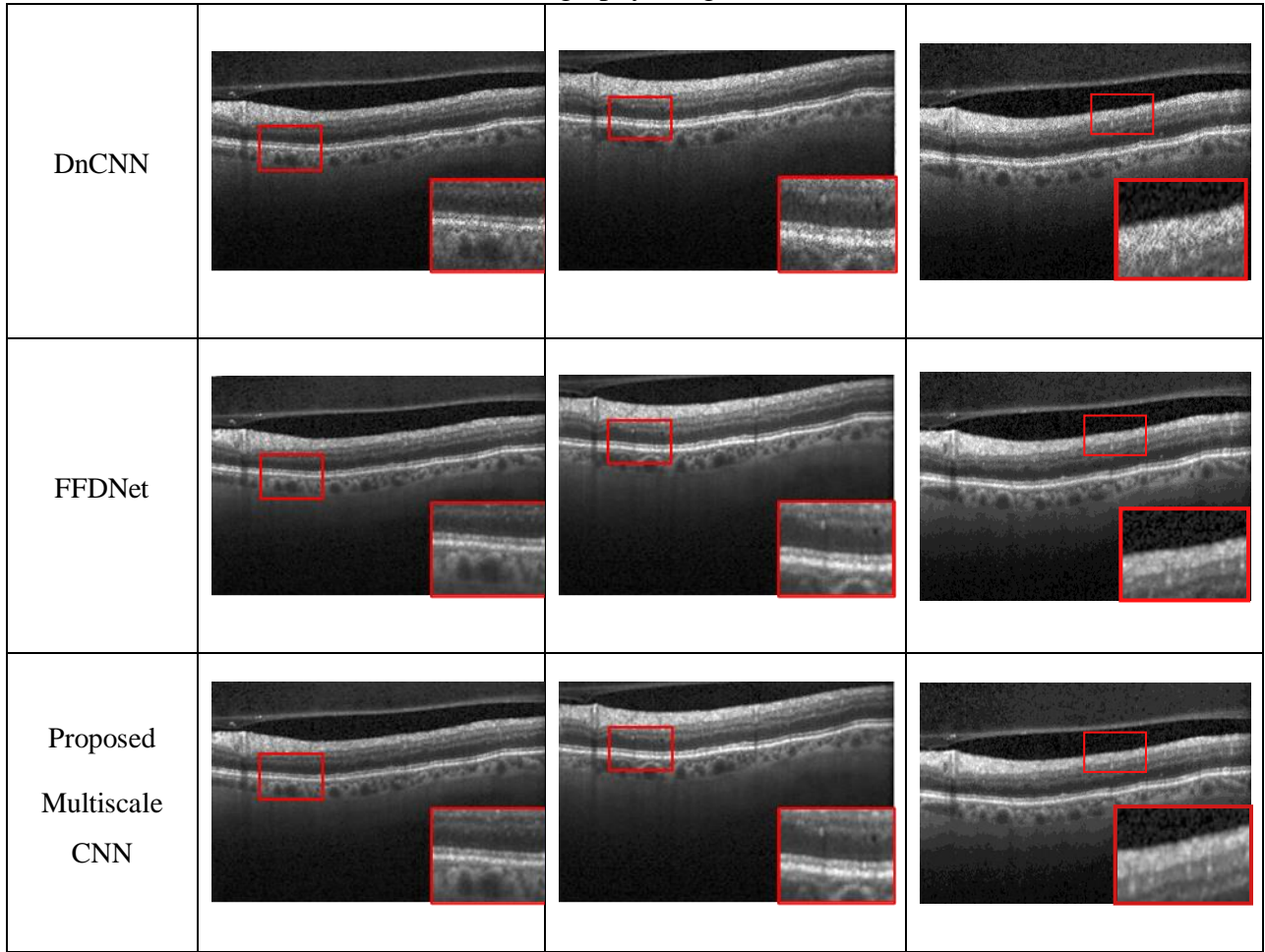|			
$\sigma=25$	30.1192	32.4322	**33.3787**
$\sigma=30$	29.5087	32.1706	**33.2912**
$\sigma=35$	28.753	32.063	**33.2711**
$\sigma=40$	28.4559	31.2771	**32.3763**
$\sigma=45$	28.2737	30.0654	**30.8677**
$\sigma=50$	28.2617	28.484	**30.4285**
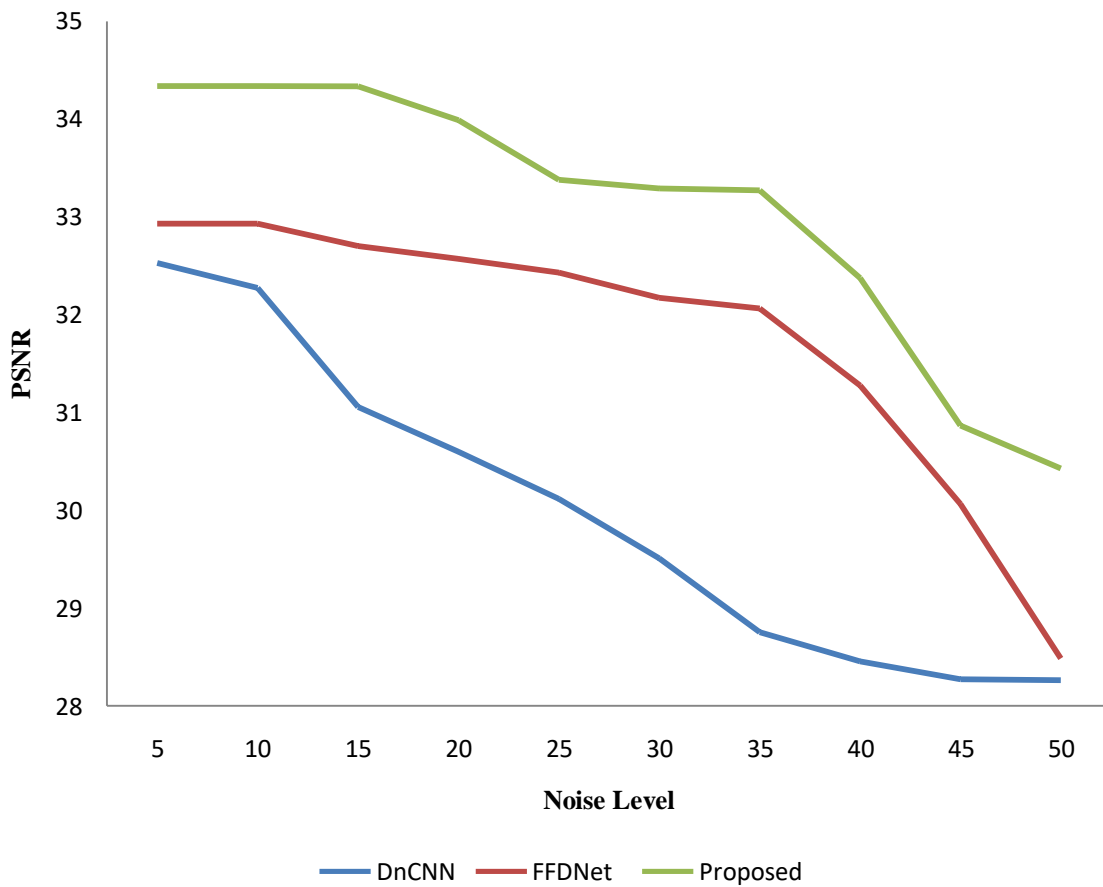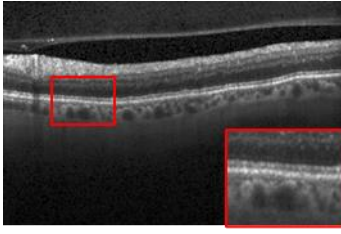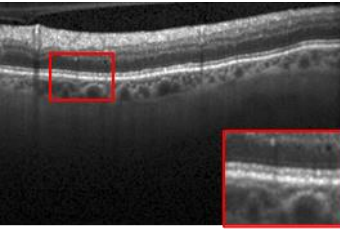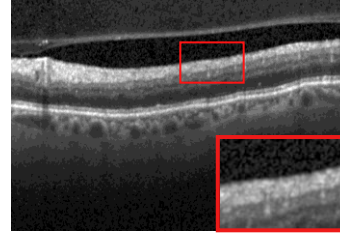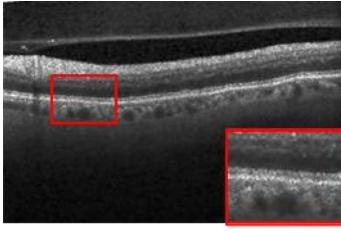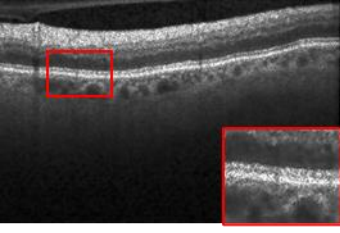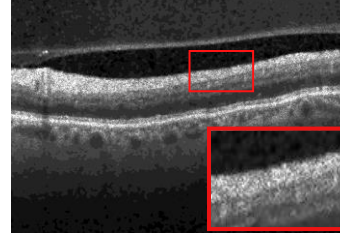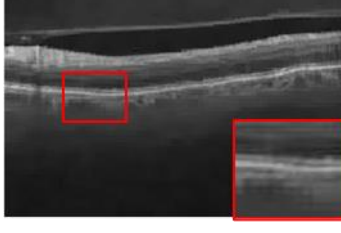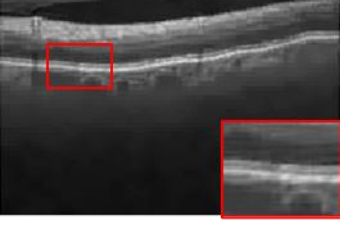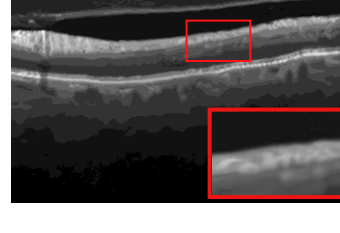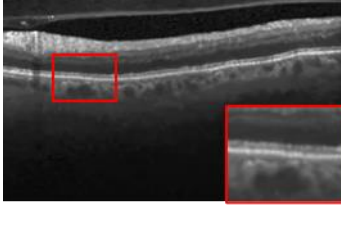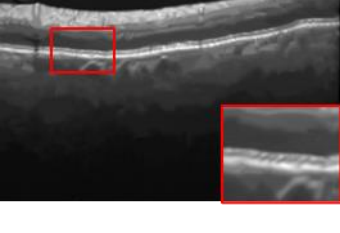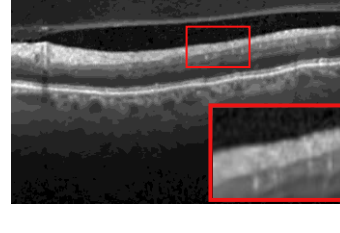
Figure 3.PSNR values for denoising models at different noise level

Table 5. Qualitative results of different denoising models for three OCT images with noise level 25
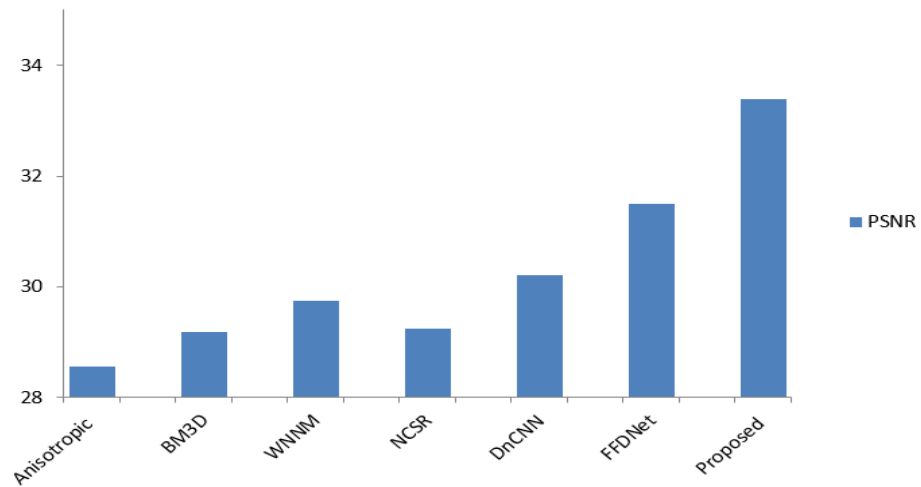
Figure 4. Performance of Proposed model with state-of-the-art methods in terms of PSNR
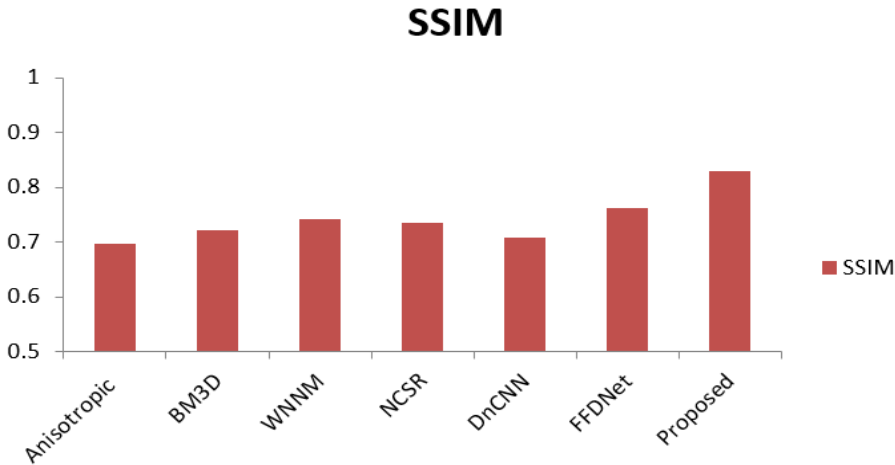
SSIM

Figure 5. Performance of Proposed model with state-of-the-art methods in terms of SSIM

NK

Figure 6. Performance of Proposed model with state-of-the-art methods in terms of NK

NAE

Figure 7. Performance of Proposed model with state-of-the-art methods in terms of NAE

LMSE

Figure 8. Performance of Proposed model with state-of-the-art methods in terms of LMSE

AD

Figure 9. Performance of Proposed model with state-of-the-art methods in terms of AD
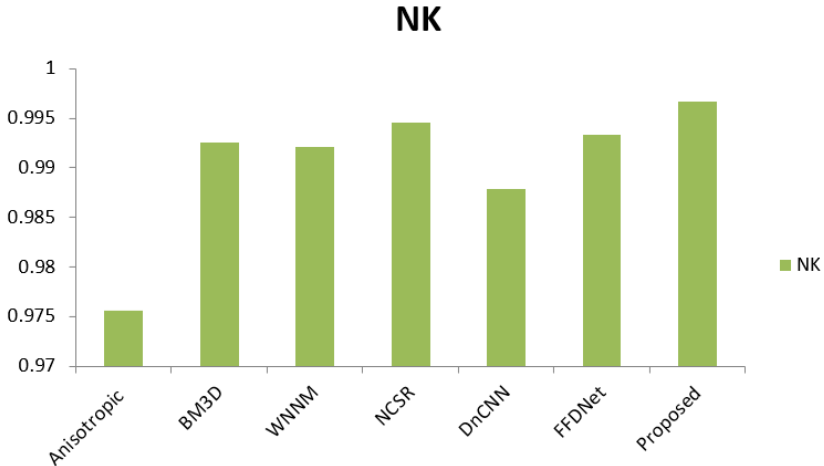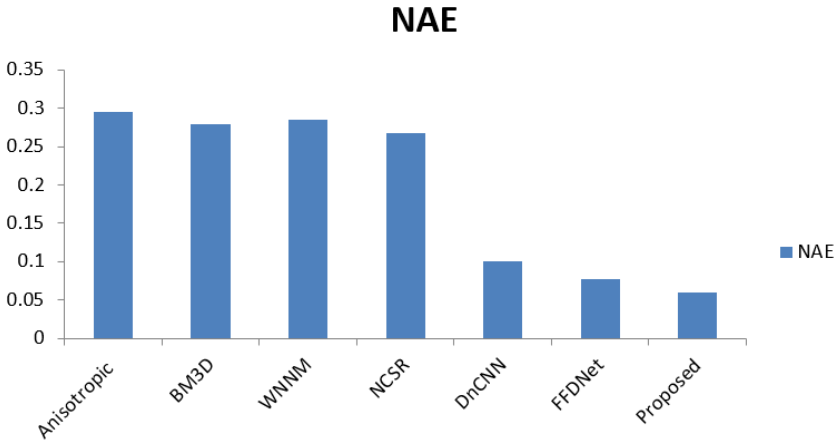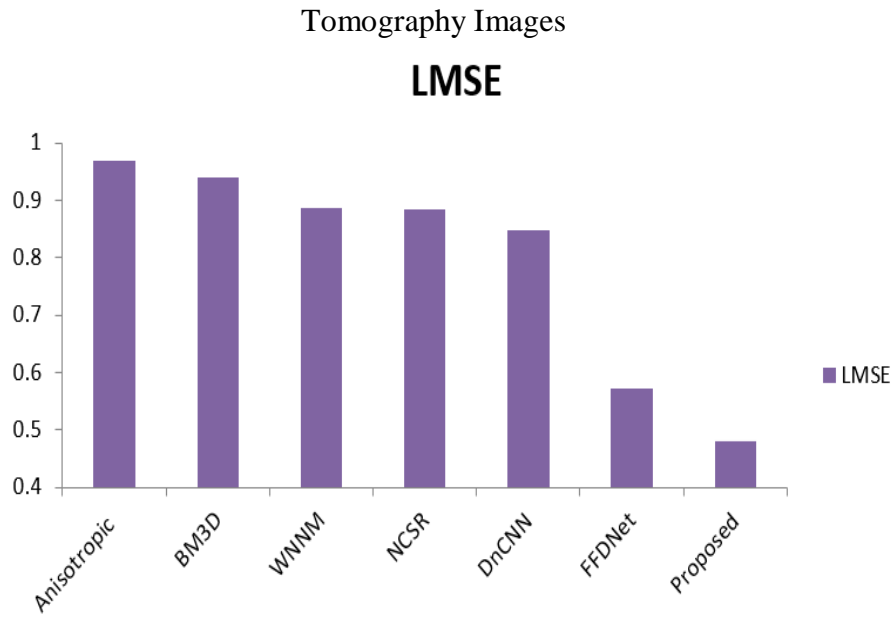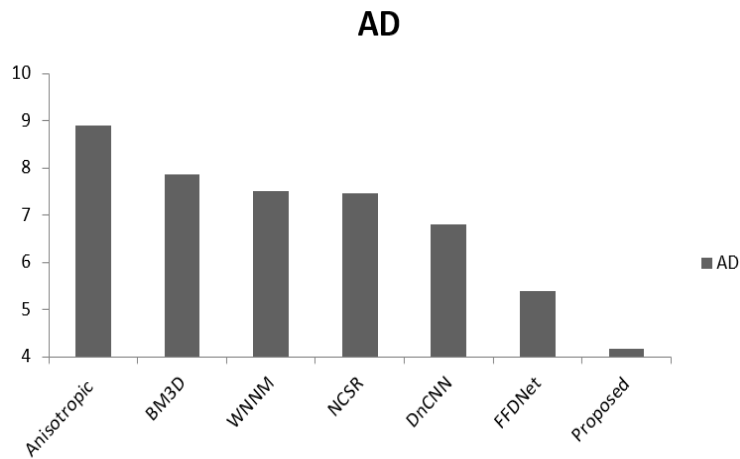
4.5.Results and Discussion

Table 3 illustrates the quantitative analysis of state-of-the-art and proposed model with different noise level 15,25,35,50 and 75.The results reveal that the proposed model achieves better PSNR value than the others for all the noise values. It maintains a slight slope of PSNR value until noise level 50. For higher noise level, PSNR value starts decreasing to a considerable value. Similarly, while Comparing to the FFDNet, the proposed model shows a significant difference in PSNR and SSIM values for lower to moderate noise level 15,25,35 and 50.

When huge noises are added, the differences in PSNR value between FFDNet and the proposed model are minimum. Table 4 compares PSNR values of state-of-the-art with proposed model for noise level σ =5 to σ =50.Table 5 shows the Qualitative results of different denoising models for three OCT images with noise level 25.

Figure 3 depicts the variations in PSNR value based on noise levels for three different deep learning models namely DnCNN, FFDNet and proposed CNN shows the visual representation of the competing methods. It is observed that anisotropic diffusion method results in severe noisy presence. The model BM3D, WNNM and NCSR performs a better denoising but it smudges the image. DnCNN model over smooth the images.

The proposed model offers a better visualization results than the other models. Figures 4,5,6,7,8 and 9 compares the performance of proposed model with other competing methods in terms of metrics like PSNR, SSIM, NK, NAE, LMSE and AD respectively.

4.6. Running Time

Table 6 demonstrates the running time of deep learning models for OCT images with different patch sizes 128 x 128,256 x 256 and 512 x 512.The experiment is performed with noise level 25 in a system with Intel® Core™ i5-8300H CPU @ 2.30GHz, 8 GB RAM and Nvidia GeForce GTX. It exposes that the proposed model has a fast-running time than the other models.

Table 6. Running time of Deep learning model for different patch sizes 128 x 128, 256 x 256 & 512 x 512

Running time (in seconds)		
128x128	256x256	512x512
DnCNN		
6.62	6.59	6.78
FFDNet		
4.95	4.88	4.89
Proposed Multiscale CNN		
4.88	**4.86**	**4.86**

A Multiscale Convolutional Neural Network for Speckle Noise Removal in Optical Coherence Tomography Images

5. Conclusion

This paper proposed a Multiscale CNN architecture endeavoring speckle noise removal in OCT images. The training is carried on gray scale images with added speckle noise. Duke dataset is used to illustrate the experiment of proposed model. The model works on downsampling images with average network depth. The Multiscale convolutions along with the batch normalization significantly improve the denoising performance. The experimental result shows that the proposed architecture outperforms the state-of-the-art methods in different parametric metrics. In the future work, the proposed model will be involved in the segmentation of retinal layers and classification of OCT images to identify retinal diseases.

References

1. M. Schmitt, "Optical coherence tomography (OCT): a review," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 5, no. 4, pp. 1205-1215, July-Aug. 1999, doi: 10.1109/2944.796348.

2. Bhende M, Shetty S, Parthasarathy MK, Ramya S. Optical coherence tomography: A guide to interpretation of common macular diseases [published correction appears in Indian J Ophthalmol. 2018 Mar;66(3):485]. *Indian J Ophthalmol*. 2018; 66(1): 20-35. doi:10.4103/ijo.IJO_902_17.

3. Fujimoto, J. G., Brezinski, M. E., Tearney, G. J., Boppart, S. A., Bouma, B., Hee, M. R., & Swanson, E. A. (1995). Optical biopsy and imaging using optical coherence tomography. *Nature medicine*, *1*(9), 970-972

4. Schmitt, J. M., Xiang, S. H., & Yung, K. M. (1999). Speckle in optical coherence tomography.

5. J. W. Goodman, "Some fundamental properties of speckle," J. Opt. Soc. Am. 66, 1145-1150 (1976).

6. Pircher M, G¨otzinger E, Leitgeb RA, Fercher AF, Hitzenberger CK (2003) Speckle reduction in optical coherence tomography by frequency compounding. J Biomed Opt 8(3): 565–570.

7. Avanaki MR, Cernat R, Tadrous PJ, Tatla T, Podoleanu AG, Hojjatoleslami SA (2013) Spatial compounding algorithm for speckle reduction of dynamic focus oct images. IEEE Photon TechnolLett 25(15): 1439–1442.

8. Wang H, Rollins AM (2009) Speckle reduction in optical coherence tomography using angular compounding by b-scan doppler-shift encoding. J Biomed Opt 14(3): 030512.

9. Pircher M, Hitzenberger CK, Schmidt-Erfurth U (2011) Polarization sensitive optical coherence tomography in the human eye. Prog Retin Eye Res 30(6): 431–451.

10. P. Perona and J. Malik, Scale-Space and Edge Detection Using Anisotropic Diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7): 629-639, July 1990

11. A. Buades, B. Coll and J. M. Morel, "A non-local algorithm for image denoising", *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 60-65, Jun. 2005.

12. K. Zhang, W. Zuo, S. Gu and L. Zhang, "Learning deep CNN denoiser prior for image restoration", *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2808-2817, Jun. 2017.

13. H. C. Burger, C. J. Schuler and S. Harmeling, "Image denoising: Can plain neural networks compete with bm3d?", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2392-2399, 2012.

14. G. Li and Y. Yu, "Visual Saliency Detection Based on Multiscale Deep CNN Features," in *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5012-5024, Nov 2016, doi: 10.1109/TIP.2016.2602079.

15. Z. Gong, P. Zhong, Y. Yu, W. Hu and S. Li, "A CNN With Multiscale Convolution and Diversified Metric for Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3599-3618, June 2019, doi: 10.1109/TGRS.2018.2886022.

16. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142–3155, July 2017.

17. Gour, N., Khanna, P. Speckle denoising in optical coherence tomography images using residual deep convolutional neural network. *Multimed. Tools Appl* 79, 15679–15695 (2020). https://doi.org/10.1007/s11042-019-07999-y.

18. K. Zhang, W. Zuo and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising," in *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608-4622, Sept. 2018, doi: 10.1109/TIP.2018.2839891.

19. Anoop, B. N., Kalmady, K. S., Udathu, A., Siddharth, V., Girish, G. N., Kothari, A. R., & Rajan, J. (2021). A cascaded convolutional neural network architecture for despeckling OCT images. *Biomedical Signal Processing and Control*, *66*, 102463.

20. Zhengjie Shen, Manhui Xi, Chen Tang, Min Xu, and Zhenkun Lei, "Double-path parallel convolutional neural network for removing speckle noise in different types of OCT images," Appl. Opt. 60, 4345-4355 (2021).

21. Li, D., Yu, W., Wang, K., Jiang, D., &Jin, Q. (2021). Speckle noise removal based on structural convolutional neural networks with feature fusion for medical image. Signal Processing: Image Communication, 99, 116500.

22. P. P. Srinivasan, L.A. Kim, P.S. Mettu, S.W. Cousins, G.M. Comer, J.A. Izatt, and S. Farsiu,"Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images", Bio Medical Optics Express, 5(10), pp. 3568-3577, 2014.

23. D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in International Conference for Learning Representations, 2015.

24. Dong, W., Zhang, L., Shi, G., & Li, X. (2012). Nonlocally centralized sparse representation for image restoration. *IEEE transactions on Image Processing*, *22*(4), 1620-1630.

25. Dabov, K., Foi, A., Katkovnik, V., &Egiazarian, K. (2006, February). Image denoising with block-matching and 3D filtering. In *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning* (Vol. 6064, p. 606414). International Society for Optics and Photonics.

26. Gu, S., Xie, Q., Meng, D., Zuo, W., Feng, X., & Zhang, L. (2017). Weighted nuclear norm minimization and its applications to low level vision. International journal of computer vision, 121(2), 183-208.

27. Eskicioglu, A. M., & Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Transactions on communications*, *43*(12), 2959-2965.

28. Mrak, S. G. M. G. (2004). Reliability of objective picture quality measures. *Journal of Electrical Engineering*, *55*(1-2), 3-10.

29. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4): 600–612.

30. Hore A, Ziou D (2010) Image quality metrics: PSNR VS. SSIM. In: 2010 20th international conference on pattern recognition (ICPR), IEEE, pp. 2366–2369.