

Received 4 June 2023, accepted 20 June 2023, date of publication 23 June 2023, date of current version 5 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3289162

RESEARCH ARTICLE

Effective Deep Learning Approach to Denoise Optical Coherence Tomography Images Using BM3D-Based Preprocessing of the Training Data Including Both Healthy and Pathological Cases

PIOTR BOGACKI¹ AND ANDRZEJ DZIECH

Faculty of Computer Science, Electronics and Telecommunications, Institute of Telecommunications, AGH University of Science and Technology, 30-059 Kraków, Poland

Corresponding author: Piotr Bogacki (pbogacki@agh.edu.pl)

This work was supported by Narodowe Centrum Bada i Rozwoju under Grant Agreement POIR.01.01-00-0397 / 19-00.

ABSTRACT Retinal diseases are significant cause of visual impairment globally. In the worst case they may lead to severe vision loss or blindness. Accurate diagnosis is a key factor in the right treatment planning that can stop or slow the disease. The examination that can aid in the right diagnosis is Optical Coherence Tomography (OCT). OCT scans are susceptible to various noise effects which deteriorate their quality and as a result may impede the analysis of their content. In this paper, we propose a novel and effective method for OCT image denoising using a deep learning model trained on pairs of noisy and clean scans obtained by BM3D filtering. A comprehensive dataset of 21926 OCT scans, collected from 869 patients (1639 eyes), covering both healthy and pathological cases, was used for training and testing of the proposed scheme. The method was validated taking into account quantitative metrics concerning image quality. In addition, the proposed denoising scheme was evaluated by analyzing the impact of applying it in the eye disease classification based on Convolutional Neural Networks (CNNs) where we obtained the improvement of around 1-3 pp (percentage point). A separate dataset of 25697 scans collected from 1910 patients (2953 eyes) was used for this purpose. The conducted experiments have proved that the method can be applied as a preprocessing step in order to provide better disease classification results and can be useful in other OCT image analysis tasks. The proposed solution is much faster and perform better than the classical BM3D filter (over ninetyfold speed-up) and other related methods, especially when a big set of images needs to be processed at once. Furthermore, the use of the diverse dataset show the benefit over methods which are based on using only healthy scans for the training of the neural network.

INDEX TERMS Optical coherence tomography, image denoising, medical applications, speckle noise removal.

I. INTRODUCTION

Optical Coherence Tomography (OCT) has become an indispensable clinical tool in ophthalmology [1], [2], neurology [3], cardiology [4], [5] and dermatology [6]. It is a well-established medical imaging technique that generates high-resolution, cross-sectional images of anatomical structure. The procedure enables three-dimensional, non-invasive,

The associate editor coordinating the review of this manuscript and approving it for publication was Ikramullah Lali.

in vivo and real time visualization of architectural morphology of the biological tissue. For instance, in ophthalmology, by observing retina, retinal nerve fiber layer (RNFL) and optic nerve head (ONH), it is possible to enhance early diagnosis and staging of ocular diseases and monitor the efficacy of a treatment [1], [2], [7].

The OCT images are obtained by measuring the interference between the backscattered infrared light from the biological sample and the light reflected from a reference mirror [1], [8]. That is why, the method is susceptible to

the presence of coherent noise (speckle noise) which has a granular pattern [9]. Another type of noise which is present in OCT scans is a system noise. It is a variation of the measured signal caused by some properties of the light detection mechanism, such as shot, relative intensity or thermal noise from the detector [10]. These phenomena are inevitable and result in degraded quality of OCT scans, reduced image contrast and impaired diagnostic interpretation of OCT data [11], [12], [13], [14]. This, in turn, may lead to the limitation of accurate layer segmentation due to imprecise delineation of the intra-retinal layers [15], [16] and deterioration of retinal thickness measurements [17], [18]. For instance, wrong estimation of RNFL thickness can affect the correct diagnosis of glaucoma [19].

In order to enhance the image quality, various image denoising methods can be used. We can differentiate both hardware and software approaches. Hardware techniques reduce noise through frequency compounding [20], [21] or spatial compounding which is based on multi-frame averaging [22], [23], [24], [25]. Although the latter method improves the contrast and image quality, it is prone to registration errors that cause edge blurring [26]. Besides, it requires longer acquisition time which may entail further problems especially with older patients that can struggle with remaining fixated during the examination [27]. On the other hand, software methods can refine the image quality through numerical operations [12], [14], [28], [29], [30], [31], [32], [33], [34], [35], [36] or filtering [37], [38], [39], [40], [41], [42], [43], [44], [45], [46]. These algorithms are realized in the post-processing step after the scans are acquired. Therefore, they are not affected by the aforementioned problems but they may be limited by the higher computational complexity and the need for appropriate parameter selection.

The solution to all the above-mentioned issues might be the use of deep learning. Although it requires a solid dataset and some specified time to train the suitable model, it can bring satisfying benefits at a later stage. Deep learning has been applied in many medical applications, including these related to OCT. For instance, it is widely exploited in layer segmentation [15], [47], [48], [49], [50], [51], [52], [53], classification [54], [55], [56], [57] and denoising [58], [59], [60], [61], [62], [63], [64], [65] tasks.

In case of OCT image denoising, deep learning methods usually require a proper training dataset that contains pairs of clean and noisy images. There are various approaches to generate such pairs. For example, frame averaging or adding artificial noise to the clean samples might be performed. However, the former method requires a large number of repeated input images to compute the ground truth. In turn, an inherent limitation of the latter approach is that the generated images may not accurately represent the statistical properties of noise in real images. Therefore, the trained models may not be effective when applied to real scans.

Another important point is the selection and number of patients involved in the study. Obviously, the greater the quantity and diversity of the scans are, especially in training

phase, the better the representation and generalization of the dataset are. Below, we discuss these aspects in some related works.

For instance, in [58] the clean images are obtained by performing multi-frame averaging and the noisy images are created by adding an artificial noise with the right distribution. A total of 20 healthy subjects were recruited for the study and in total 3880 scans were obtained.

Another option for the creation of clean/noisy pairs, presented in [59], is to select one frame from a series and treat it as a noisy sample and then perform multi-frame averaging on the whole series of scans and make the resulting image a corresponding clean sample. For data acquisition, 2350 scans from 47 healthy eyes were obtained.

In [60] except for signal averaging also contrast enhancement is performed in order to generate clean scans. For training 512 scans of healthy eyes were used while for testing 36 scans coming from both healthy and pathological eyes were selected.

The method described in [61] is also based on frame averaging. Six OCT volumes were acquired from both eyes of 38 healthy patients. These scans were then registered and averaged to create the ground truth denoised image. Due to inaccurate registration only 69 volumes (each consisting of 180 B-scans) from healthy eyes were used for further training and analysis. As a result, 74520 clean/noisy scan pairs were produced. In addition, 1080 B-scans from patients with glaucoma were acquired but they were only used to evaluate the method.

In [63] volumetric OCT datasets of healthy human retina centered on the fovea and optic nerve head were acquired to train and test the self-fusion neural network. Each volume contained 2500 raw B-scans (500 sets of 5-repeated frames). The repeated frames were averaged and the resulting frames were self-fused together with 6 adjacent frames (3 before and 3 after the current frame) in order to achieve ground truth images of high quality for the neural network training. The model was trained on 9 ONH volumes and validated on 3 fovea volumes. For testing, 3 ONH and 3 fovea volumes were used. Additionally, OCT images from external datasets were added as test images. This group also contained the samples with pathological features present.

Another approach was presented in [64] where the authors do not need the pairs of clean and noisy samples but the pairs of scans from the same location but with a different noise level. In total, images from 16 subjects, resulting in 3895 pairs of B-scans at different locations of the retina, were included for training. For the evaluation 3 pairs of single and 128x averaged scans acquired at ONH from distinct subjects were used.

Similarly as in the previous work, in [65] the authors propose a deep learning method for noise reduction in OCT images without the need for obtaining noise-free ground truth as labels. In contrast, they need pairs of B-scans from the same sample location. Then, one noisy image is taken as the input while another noisy image is treated as the label.

The dataset consists of 15 subsets of OCT images of various sample materials, including finger nails, hand palms, tomato, sample tooth, plastic tubes and thin films. In total, 3750 OCT scans were collected in this study.

It is also relevant to select the right method for introducing the artificial noise to the images. Some approaches assume that the noise has additive form [58], [66] and other apply a multiplicative noise model [59], [67]. Another key factor here is the distribution of the noise. Some methods in the literature approximate the speckle noise distribution to Gaussian [58], [66], [67], Rayleigh [68], [69]/Generalized Gamma [70] or Poisson [46]. However, each model type has its own pros and cons and it is hard to find a universal solution.

The denoising scheme presented in this paper assumes using two approaches for the generation of pairs of clean and noisy scans. The first method is the standard one that simply adds a Gaussian noise to the original clean scans and it is used for comparison purpose. The second approach applies the BM3D filter to the noisy images selected from the original dataset. The main characteristic of this approach is the fact that the noisy images represent the actual noise which is present in OCT scans and not the one which is artificially added. Another aspect is that the clean images are obtained by using BM3D algorithm which is commonly used for image denoising. This way, we will try to train the network so that it tries to reproduce the results obtained by using the BM3D filter. Similar attempt was proposed in [71] where the structure of the convolutional neural network is designed in such a way that it mimics some steps from the computational pipeline of BM3D method, e.g. its block matching stage, by adding extraction and aggregation layers in the model. In contrast, our method is based on the preprocessing of the training dataset so that the network can try to replicate the filtering of BM3D algorithm.

One of the main advantages of the proposed method in contrast to most of methods described in the literature consists in using the dataset that contains both the scans from healthy patients and the ones with pathological features present in the image. Moreover, it consists of 21926 samples acquired from 869 patients (1639 eyes) which guarantees a good representation and diversity of potential cases we may face during OCT examinations. Besides, the proposed method is validated taking into account not only quantitative metrics concerning image quality but, in addition, it is evaluated by analyzing the impact of applying it in the eye disease classification. For this purpose, a separate dataset of 25697 scans collected from 1910 patients (2953 eyes) was used. Although the quantitative metrics can give an overview about the general effect on image quality improvement, they may be easily tricked. Therefore, it is of great importance to assess the performance of the given method by examining how it behaves in some specific OCT image analysis tasks.

The paper is organized as follows. The following section is dedicated to the methodology of the proposed denoising scheme. Section III contains experimental results and

their discussion. Finally, section IV presents conclusions and future work.

II. METHODOLOGY

A. USED DATASET

The dataset used in this study consists of 21926 scans. They were collected during the realization of the INDOK project [72] that aims to develop a system for automatic diagnosis of retinal diseases based on OCT images with the use of artificial intelligence. The dataset contains the samples from both healthy and sick eyes. In total, the scans from 869 patients were collected. Among the pathological changes present in the 'sick' scans in total 19 features were identified, e.g. cyst, drusen, hyperreflective foci (HRF), epiretinal membrane (ERM), retinal detachment (RD), macular hole (MH) and macular pseudo hole (MpH), serous and fibrovascular pigment epithelial detachment (sPED/fvPED). The collected scans are of various quality - some are clean and other have significant amount of speckle and/or system noise visible. All scans have been precisely tagged with proper annotations by professionals from the ophthalmology field. The method for splitting the dataset into two subsets regarding the quality of the scans was described in detail in the following subsection. The exact representations for healthy/sick and clean/noisy scans are presented in Table 1.

TABLE 1. Dataset representation for healthy/sick and clean/noisy scans.

	clean	noisy	Total
healthy	3094	3125	6219
sick	6267	9440	15707
Total	9361	12565	

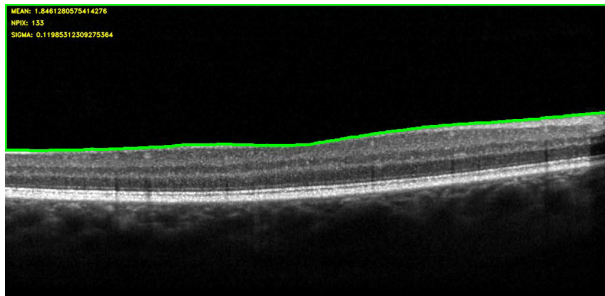
B. SPLIT INTO CLEAN AND NOISY SAMPLES

In order to divide the samples into either a clean or a noisy subset, it was noticed that it is much easier to detect system noise in the upper part of the scan which is located above the retinal layers. Therefore, a few metrics have been selected in order to describe this region. Additionally, a whole image is also analyzed in order to estimate the level of speckle noise. We selected these metrics based on our observation concerning the appearance and a character of noise which is present in OCT scans. All these metrics are used to group the scans into one of the mentioned categories. These metrics include:

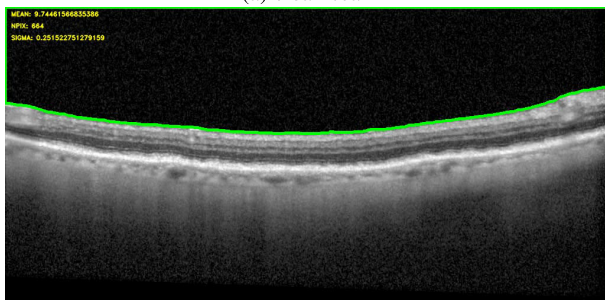
- *mean* - mean value of pixels belonging to the upper region of the image
- *npix* - number of non-zero pixels belonging to the upper region of the image
- *npix_coef* - relation between *npix* and the area of the upper region of the image
- *sigma* - noise variance estimation measure presented in [73], calculated for the upper part of the scan
- *sigma_global* - noise variance estimation measure, calculated for the whole image

TABLE 2. Results of noise level estimation metrics for both clean and noisy scans.

	<i>mean</i>	<i>npix</i>	<i>npix_coef</i>	<i>sigma</i>	<i>sigma_global</i>
clean	3.87 ± 1.60	658.84 ± 833.53	0.0049 ± 0.0053	0.258 ± 0.180	1.206 ± 0.657
noisy	14.98 ± 9.69	12285.67 ± 22774.97	0.0744 ± 0.1184	4.903 ± 7.926	3.208 ± 2.149



(a) clean scan



(b) noisy scan

FIGURE 1. Sample scans with outlined top region of interest.

The upper regions in the scans were detected through Gaussian blurring, image thresholding, some morphological operations and analysis of the location and area of the identified blobs. The sample difference between clean and noisy images with outlined upper regions of scans is shown in Fig. 1.

The distribution of values for each measure is depicted in corresponding histograms in Figs. 2a to 2e and in Table 2.

As can be seen in the presented results, none of the metrics allows for easy separation of data and it is difficult to define a specific threshold for any of them. That is why, it was necessary to apply a more complex solution. Hence, it was decided to select 500 random samples for both classes (clean and noisy) which were consulted with an ophthalmologist and then train an SVM model based on these annotations and aforementioned metrics. 80% of samples from each group were used for training and the remaining 20% for test. The training of the model was realized in Python with the use of *scikit-learn* package [74]. The trained SVM model achieved the accuracy of 95%. As a result, 9361 clean samples and 12565 noisy samples have been obtained.

C. NETWORK ARCHITECTURE

The network architecture selected for the task of OCT scans denoising is a modified U-Net architecture introduced for the first time in 2015 in [75]. It was originally used for biomedical

image segmentation. However, due to its characteristics and after slight modifications it can be also applied for OCT image denoising. The U-Net network consists of two main blocks, i.e. downsampling and upsampling ones. The downsampling block (or encoder) performs a specific number of convolutions so that the input image is downsampled to a feature map. By contrast, the upsampling block (or decoder) through transposed convolution operations upsamples this feature map back to the original input image size. In addition, the so-called skip connections help to avoid losing relevant information from the encoder part and preserve the spatial relationships. It is done by transferring information from every downsampling layer in the encoder to their corresponding upsampling layer in the decoder. In this way, the network learns how to retain the knowledge about both local (textures) and contextual (spatial arrangement) characteristics of a tissue [58].

In our study, the network takes as the input the images of size 256×192 pixels. It performs four downsampling and corresponding upsampling operations. It has over 8 million trainable parameters. Adam optimizer is used for training with default learning rate equal to 0.001 and the mean squared error (MSE) is selected as a loss function. The number of epochs is 40 and the batch size is set to 32. The trainings were held on NVIDIA GeForce RTX 2800 Ti with CUDA v10.1 and cuDNN v7.6.5 acceleration.

D. TRAINING AND TESTING OF THE NETWORK

In this study, two approaches for the generation of pairs of clean and noisy scans have been selected. The first approach involves adding a Gaussian noise to the original clean samples. 80% of the pairs created this way are used for training. For testing, 20% of the remaining pairs were assigned. In contrast, the second approach aims at emulating the BM3D filter behavior. BM3D filtering is a very common method used for OCT image denoising [38], [70]. In order to let the network reduce the noise in a similar way, the noisy dataset was subjected to BM3D filtering implemented in Python [76]. This time, similarly as in the first approach, the resulting pairs of noisy and clean samples were split into training and testing subsets using the same ratio as previously (80:20).

Another option that was examined consisted in taking this whole noisy subset, together with the corresponding scans subjected to BM3D filtering, for training. In turn, the pairs of clean samples together with the samples with added Gaussian noise were utilized for testing. Although this leads to the situation when the training and testing samples come from different data distributions and have a slightly different nature, it was still worth verifying how the network could behave.

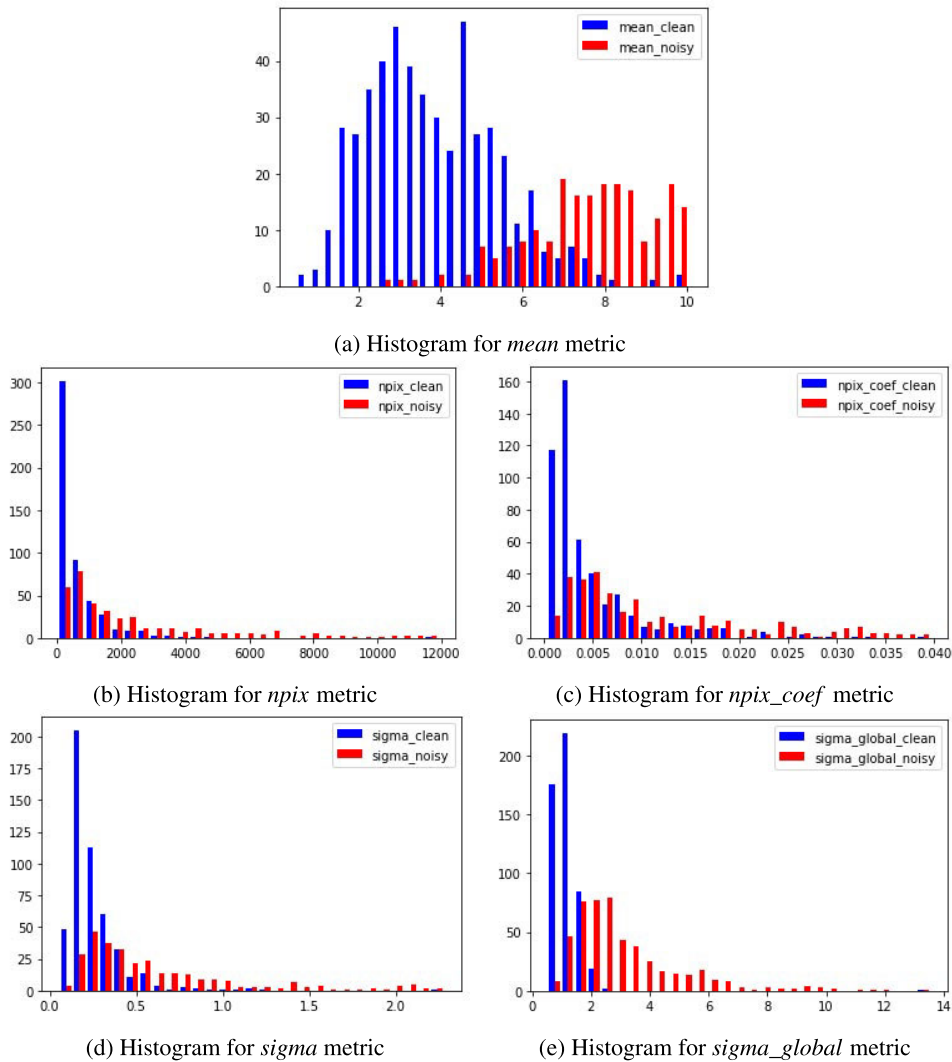


FIGURE 2. Distribution of values for noise level estimation metrics.

When applying BM3D filtering, the algorithm takes as the input the value of sigma parameter which is related to the expected noise level. In our experiments three variants have been taken into account:

- 1) σ parameter is assigned based on the value calculated following the method for noise variance estimation presented in [73] [SIGMA_REAL]
- 2) σ parameter is the previous value incremented by 10 [SIGMA_PLUS]
- 3) σ parameter is equal to 20 [SIGMA_20]

In case of adding a Gaussian noise to the samples from the clean subset, the following ranges for σ parameter are taken into account:

- 1) $\sigma \in < 15, 30 >$ [ADDNOISE_15_30]
- 2) $\sigma \in < 5, 30 >$ [ADDNOISE_5_30]
- 3) $\sigma \in < 25, 30 >$ [ADDNOISE_25_30]

The limits for these three ranges were selected based on potential values of noise variance in OCT scans and in order to examine various scopes.

All the above-mentioned variants are presented in Fig. 3.

E. QUANTITATIVE MODEL EVALUATION

In order to assess the performance of all denoising approaches, the following image quality metrics have been selected:

- Mean Squared Error (MSE) - it represents the cumulative squared error between two images. It can be calculated using the below formula (1):

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_1(m, n) - I_2(m, n)]^2 \quad (1)$$

where: M - the width of the image
 N - the height of the image

- Peak Signal to Noise Ratio (PSNR) - it measures the peak error between two images. It can be expressed by the following formula (2):

$$PSNR = 20 \cdot \log_{10} \left(\frac{MAX}{\sqrt{MSE}} \right) \quad (2)$$

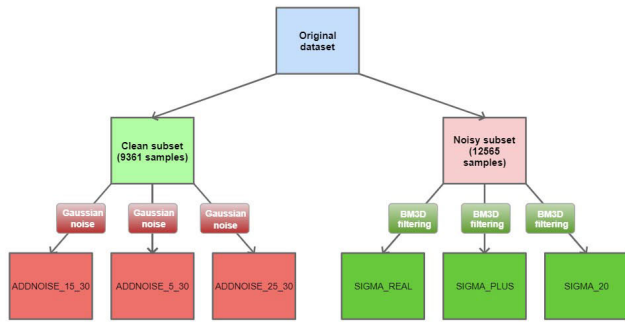


FIGURE 3. Dataset variants after BM3D denoising and adding of Gaussian noise.

where: MAX - the maximum possible pixel value of the image

- Structural Similarity (SSIM) - it is a quality assessment metric based on the visual degradation of structural information (changes in local structure and contrast between two images). It can provide a good approximation to human visual perception [77]. The parameter value can be obtained using the below formula (3):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

where: μ_x, μ_y - mean intensities of two images being compared
 σ_x, σ_y - standard deviations of two images
 σ_{xy} - cross-covariance between both images
 C_1, C_2 - two constants to stabilize the division

- Multi-Scale Structural Similarity (MS-SSIM) - it is based on multiple SSIM measurements at different image scales [78]. It is calculated with the use of the following formula (4):

$$MS-SSIM(x, y) = \prod_{i=1}^M SSIM(x_i, y_i) \quad (4)$$

where: x_i, y_i - the images at i^{th} scale
 M - the number of scales

A small MSE indicates minor error. The larger a PSNR value is, the better quality an image has. Both SSIM and MS-SSIM return values from the range $< 0, 1 >$, where 0 indicates no similarity and 1 implies that both images are identical.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. DENOISING PERFORMANCE

At the beginning, we use cross-validation to evaluate the efficiency of the deep learning approaches. Therefore, both datasets with pairs of clean and noisy scans are divided into training and testing subsets with the ratio 80:20. The results for the networks trained on both datasets described in the subsection II-D are presented in Tables 3 and 4.

As might be observed, both denoising schemes seem to work and in most cases provide improved results taking into account particular metrics. It is worth noting that in case of SIGMA_REAL approach the metrics for the noisy scans seem to be better. However, this is caused by the little effect introduced by denoising with sigma parameter which is equal to the value obtained by using the noise variance estimation formula [73]. As a result, noisy and clean (after BM3D denoising) images are very similar and the results seem to be very satisfying but in this case they should be ignored. That is why, it seems reasonable to check other variants for the sigma parameter in BM3D algorithm.

In the next step, we focus on the tests performed when the noisy subset and the corresponding samples denoised with the use of BM3D method are used for training. Then, the full clean dataset with added Gaussian noise will be treated as a testing set. Although the noise nature is different across these two datasets, such experiments can help to notice some dependencies and point out which method could be useful in various tasks.

The results for this scheme are shown in Table 5. Each test set (ADDNOISE*) is evaluated with the use of three models - UNET_REAL, UNET_PLUS and UNET_20 - trained on different training sets (SIGMA_REAL, SIGMA_PLUS and SIGMA_20, respectively). In addition to the approaches based on U-Net architecture, we also evaluate the original BM3D algorithm and BM3D-Net [71] algorithms. Moreover, we also verified how a different architecture model could behave when trained on the same two datasets (SIGMA_PLUS and SIGMA_20) as in case of U-Net. This way we are able to determine if the proposed approach is valid no matter which model is used. For this purpose RED-Net model [79] was selected. The first row is common for all seven methods since it contains the results for the metrics calculated for all the pairs of corresponding clean and noisy scans. Then, the following seven rows present the results for denoised scans from particular denoising variants.

The results obtained for the models that are trained on noisy scans and their corresponding BM3D-denoised scans (especially SIGMA_PLUS and SIGMA_20 datasets), that try to emulate the outcomes of BM3D filter, seem to be satisfying since they provide better results for the denoised scans for all the metrics in comparison to referenced noisy scans. These two schemes are the most promising in terms of applying them in other OCT image analysis tasks which is why they will be evaluated more deeply in the following subsection.

It is worth mentioning that the difference between particular approaches presented in Table 5 is statistically significant. It was verified by performing ANOVA test for all three testing scenarios and all four quality metrics. Each time the P-value measured within and between particular groups (denoising approaches) was equal to zero. This satisfies the condition that P-value is lower than 0.05 which is usually the value selected for the significance level.

TABLE 3. Denoising results for the network trained and tested on clean scans and scans with added Gaussian noise.

Dataset	Type	PSNR	SSIM	MS-SSIM	MSE
ADDNOISE_15_30	noisy	29.59 ± 0.35	0.37 ± 0.07	0.84 ± 0.04	364.48 ± 131.42
	denoised	32.29 ± 0.97	0.72 ± 0.07	0.94 ± 0.02	135.32 ± 23.33
ADDNOISE_5_30	noisy	30.43 ± 1.38	0.49 ± 0.16	0.88 ± 0.06	254.32 ± 174.82
	denoised	32.20 ± 1.01	0.71 ± 0.08	0.94 ± 0.02	138.07 ± 26.01
ADDNOISE_25_30	noisy	29.30 ± 0.22	0.30 ± 0.02	0.79 ± 0.02	520.10 ± 57.22
	denoised	32.30 ± 0.96	0.72 ± 0.07	0.94 ± 0.02	133.48 ± 22.96

TABLE 4. Denoising results for the network trained and tested on noisy scans and scans filtered via BM3D algorithm.

Dataset	Type	PSNR	SSIM	MS-SSIM	MSE
SIGMA_REAL	noisy	56.92 ± 22.16	0.96 ± 0.09	0.99 ± 0.02	28.42 ± 69.70
	denoised	29.98 ± 0.74	0.48 ± 0.09	0.86 ± 0.05	293.93 ± 236.38
SIGMA_PLUS	noisy	32.01 ± 1.72	0.78 ± 0.07	0.96 ± 0.02	83.74 ± 90.37
	denoised	32.47 ± 1.70	0.74 ± 0.13	0.93 ± 0.05	133.19 ± 165.93
SIGMA_20	noisy	31.05 ± 1.27	0.71 ± 0.06	0.95 ± 0.01	95.11 ± 49.34
	denoised	33.88 ± 2.29	0.82 ± 0.16	0.95 ± 0.06	125.35 ± 172.64

TABLE 5. Denoising results for the network trained on noisy scans and scans filtered via BM3D algorithm and tested on clean scans and corresponding scans with added noise.

Dataset (Train + Test)	Type	PSNR	SSIM	MS-SSIM	MSE
SIGMA* + ADDNOISE_15_30	noisy	29.59 ± 0.35	0.37 ± 0.07	0.84 ± 0.04	365.53 ± 131.20
BM3D	denoised	29.68 ± 0.32	0.41 ± 0.06	0.87 ± 0.03	271.78 ± 82.53
BM3D_NET	denoised	30.11 ± 0.71	0.57 ± 0.04	0.94 ± 0.01	139.23 ± 22.17
UNET_REAL	denoised	29.50 ± 0.29	0.32 ± 0.03	0.82 ± 0.02	415.20 ± 73.35
UNET_PLUS	denoised	30.05 ± 0.57	0.46 ± 0.06	0.89 ± 0.03	241.34 ± 46.71
UNET_20	denoised	30.33 ± 0.79	0.52 ± 0.07	0.91 ± 0.03	195.96 ± 46.10
REDNET_PLUS	denoised	30.76 ± 1.01	0.59 ± 0.07	0.94 ± 0.01	97.39 ± 43.61
REDNET_20	denoised	31.26 ± 1.18	0.63 ± 0.06	0.95 ± 0.01	70.69 ± 28.91
SIGMA* + ADDNOISE_5_30	noisy	30.45 ± 1.38	0.49 ± 0.16	0.88 ± 0.06	251.70 ± 174.50
BM3D	denoised	30.52 ± 1.38	0.51 ± 0.14	0.90 ± 0.04	195.03 ± 116.27
BM3D_NET	denoised	30.91 ± 1.26	0.61 ± 0.07	0.94 ± 0.01	127.03 ± 25.25
UNET_REAL	denoised	29.85 ± 0.63	0.38 ± 0.10	0.85 ± 0.05	337.12 ± 115.19
UNET_PLUS	denoised	30.76 ± 1.15	0.54 ± 0.11	0.91 ± 0.03	203.06 ± 61.88
UNET_20	denoised	31.05 ± 1.23	0.58 ± 0.11	0.93 ± 0.03	168.48 ± 51.25
REDNET_PLUS	denoised	32.09 ± 1.97	0.67 ± 0.12	0.95 ± 0.02	70.27 ± 47.84
REDNET_20	denoised	32.23 ± 1.66	0.68 ± 0.08	0.96 ± 0.01	56.00 ± 29.47
SIGMA* + ADDNOISE_25_30	noisy	29.31 ± 0.22	0.30 ± 0.02	0.79 ± 0.02	519.99 ± 57.19
BM3D	denoised	29.42 ± 0.22	0.40 ± 0.07	0.87 ± 0.04	300.38 ± 107.91
BM3D_NET	denoised	29.40 ± 0.26	0.54 ± 0.03	0.93 ± 0.01	157.02 ± 17.68
UNET_REAL	denoised	29.38 ± 0.22	0.30 ± 0.03	0.81 ± 0.02	462.57 ± 67.82
UNET_PLUS	denoised	29.55 ± 0.25	0.41 ± 0.03	0.87 ± 0.02	285.57 ± 33.23
UNET_20	denoised	29.60 ± 0.29	0.45 ± 0.04	0.89 ± 0.02	244.59 ± 30.77
REDNET_PLUS	denoised	29.74 ± 0.25	0.50 ± 0.03	0.92 ± 0.01	149.47 ± 24.76
REDNET_20	denoised	30.04 ± 0.32	0.56 ± 0.03	0.94 ± 0.01	105.05 ± 17.05

Apart from quantitative metrics analysis, it is always recommended to visually assess the outcomes of image processing algorithms. The sample groups of clean, noisy and denoised scans for UNET_PLUS and UNET_20 approaches are presented in Figs. 4a and 4b. One can see that in both noisy images there is a noticeable granular pattern present. In both denoised images it is much less visible. Besides, the level of details is higher in denoised images and they resemble the original structures from the clean scans. For instance, the difference between particular layers or some little features (like hyperreflective foci or shadows) can be much more easily perceived. This might be especially noticed in case of UNET_20 denoising.

Since quantitative metrics can be easily tricked and the visual outcomes do not guarantee how the denoising will

influence, for example, disease classification accuracy, the more elaborated evaluation is required. Such an attempt is made in the next subsection.

B. IMPACT ON DISEASE CLASSIFICATION ACCURACY

In order to evaluate the usefulness of the proposed denoising schemes, their impact on disease classification accuracy has been assessed. Therefore, three CNN-based classifiers that aim to detect the presence of three eye diseases in OCT scans were trained. These classifiers perform binary classification to one of two classes - selected disease or ‘rest’ which covers all other diseases and/or healthy scans. The diseases selected for the evaluation are Age-Related Macular Degeneration (AMD), Epiretinal Membrane (ERM) and Diabetic Retinopathy (DR). The dataset used for training and testing contains

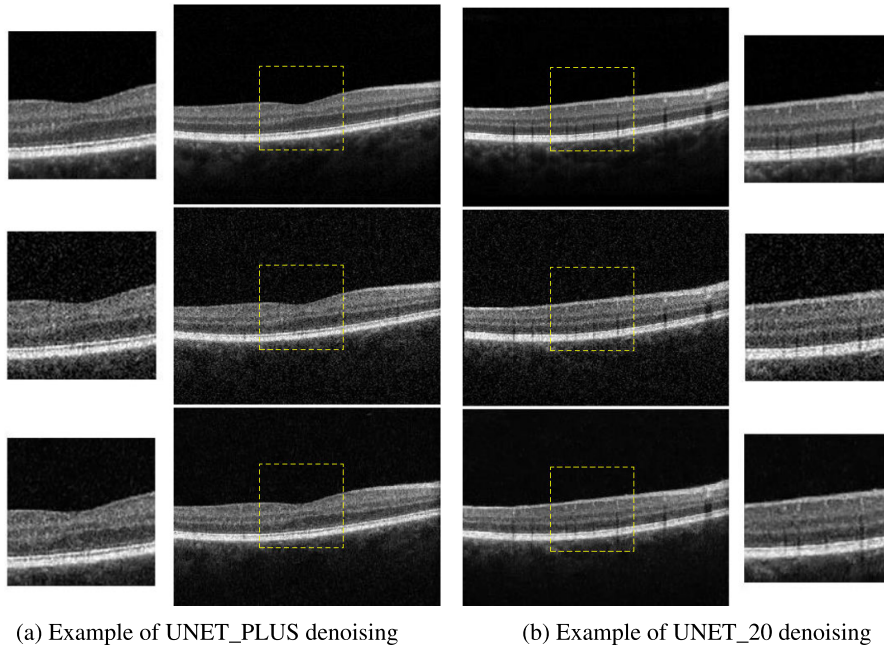


FIGURE 4. Sample groups of clean (top), noisy (middle) and denoised (bottom) scans for UNET_PLUS and UNET_20 schemes with zoomed in regions for comparison.

5539 scans with AMD present, 4704 with ERM and 711 with DR. In total 25697 scans obtained from 1910 patients and covering 2953 single eyes were used. It is worth noting that the dataset used for training and testing in disease classification task is independent of the one used for the creation of the denoising models described above. It is guaranteed by selecting the scans acquired in different time periods for both datasets. For each classifier the class balance was assured.

The exact statistics regarding the classification dataset is presented in Table 6. The datasets were split into training, validation and test subsets using a ratio 70:10:20. This split was performed in such a way that the scans of the same eye can be selected only to one subset. Some preliminary investigations indicated that the most satisfying results for AMD, ERM and DR recognition are obtained by VGG16, Resnet152 and VGG19 model architectures, respectively. The image classifiers were implemented in Python with the use of Tensorflow library [80].

For each disease, six approaches were investigated: no denoising (used as a reference for the remaining solutions), classical BM3D filtering, BM3D-Net [71] and two different architectures trained using two proposed schemes (SIGMA_PLUS and SIGMA_20), i.e. REDNET_PLUS, REDNET_20, UNET_PLUS and UNET_20, with the latter variant additionally evaluated when trained only on the healthy scans (UNET_20_HS). This way it will be possible to determine whether denoising as a preprocessing step, in general, improves the disease classification accuracy and if the proposed methods can offer a benefit over classical and other already existing approaches. What is more, in contrast to most of other similar attempts, the dataset used for training

TABLE 6. Statistics of the dataset used for training and testing of three classifiers dedicated to the recognition of AMD, ERM and DR diseases.

Dataset	Class	# scans	# patients	# eyes
AMD_vs_REST	AMD	5539	566	843
	REST	5545	1155	1676
ERM_vs_REST	ERM	4704	552	725
	REST	4714	1140	1662
DR_vs_REST	DR	711	67	92
	REST	716	417	453

of the proposed denoising models contains both healthy and pathological samples. Therefore, the impact of using such a diverse dataset is also studied by comparing last two variants (UNET_20 and UNET_20_HS).

The results for accuracy, F1 score and AUC of image classification for all three diseases are presented in Table 7. A sample ROC curve (of UNET_20 model for AMD) was shown in Fig. 5. The assignment to one of two classes (DISEASE or REST) was based on the higher probability of the given class (threshold equal to 0.5).

It can be seen from the Table 7 that the proposed denoising methods guarantee similar (in case of AMD) or better (for ERM and DR) classification performance with respect to the classical BM3D filtering. Similarly, at least one of the proposed methods may outperform the BM3D-Net. This is especially noticeable in case of ERM and DR classifier. The classifiers which were trained on the diverse datasets always prevail over their corresponding versions trained only on the healthy scans. This clearly justifies the benefit of using the extended dataset covering both healthy and non-healthy scans. Besides, all the classifiers which operate on

TABLE 7. Impact of the proposed denoising schemes on disease classification accuracy and the comparison with other denoising methods and no denoising performed.

Dataset	CNN Architecture	Denoising	Accuracy [%]	F1 score [%]	AUC
AMD	VGG16	no denoising	92.84	92.82	95.49
		BM3D denoising	93.70	93.69	96.96
		BM3D_NET denoising	93.38	93.38	97.31
		REDNET_PLUS denoising	91.22	91.18	95.55
		REDNET_20 denoising	93.52	93.51	97.06
		UNET_PLUS denoising	93.34	93.33	96.94
		UNET_20 denoising	93.65	93.65	95.73
		UNET_20_HS denoising	92.80	92.80	96.65
ERM	Resnet152	no denoising	87.90	88.02	92.20
		BM3D denoising	88.32	87.43	92.33
		BM3D_NET denoising	83.40	86.20	91.84
		REDNET_PLUS denoising	86.73	89.46	94.37
		REDNET_20 denoising	87.58	89.90	94.82
		UNET_PLUS denoising	89.11	89.08	94.90
		UNET_20 denoising	88.74	88.89	94.72
		UNET_20_HS denoising	86.31	86.58	93.56
DR	VGG19	no denoising	91.38	91.31	93.82
		BM3D denoising	91.38	91.33	94.58
		BM3D_NET denoising	92.76	92.75	93.61
		REDNET_PLUS denoising	93.10	93.09	96.82
		REDNET_20 denoising	92.41	92.39	96.07
		UNET_PLUS denoising	92.41	92.41	97.33
		UNET_20 denoising	94.48	94.48	98.83
		UNET_20_HS denoising	93.45	93.44	96.44

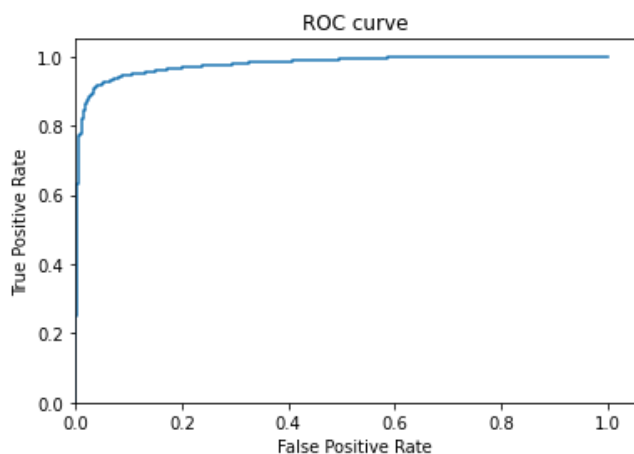


FIGURE 5. ROC curve for AMD.

the denoised images in the majority of cases can provide improved classification results. What is more, the results are better for both metrics in case of the proposed solutions in comparison to the classifiers which do not utilize denoised images. Although the improvement is of around 1-3 pp (percentage point), it is worth highlighting that the basic classifiers, with no denoising performed, could already offer the classification accuracy (or F1 score) of around 90%. In such a case making progress of every 1 pp is significant while at the same time being much more difficult to achieve. In other words, we can imagine that using one of the presented schemes can effectively lead to the situation when additional 1-3 in 100 sick patients are given a correct diagnosis and as a result can get adequate treatment on time. Eventually, the enhanced classifiers can save a considerable amount of patients by stopping or delaying their disease progression or even preventing them from getting blind.

C. DENOISING TIME

The denoising of the whole noisy subset (12565 scans) with the use of the proposed method (UNET_20 variant) took 1104.9207 s and in case of BM3D_NET it took 23161.0645 s. Due to the fact that the denoising via BM3D filtering takes much longer, only 5903 scans were processed which took 46981.46 s. The processing time for a single scan is shown in Table 8.

TABLE 8. Denoising time comparison (BM3D vs BM3D_NET vs Proposed scheme).

Method	Denoising time for a single scan [s]
BM3D	7.9589
BM3D_NET	1.8433
RED_NET	0.7252
Proposed scheme	0.0879

As might be observed, the proposed scheme offers much faster processing time, namely a ninetyfold reduction in comparison with standard BM3D approach, over twentyfold reduction in comparison with BM3D-Net and over eightfold reduction in comparison with RED-Net. All three methods were tested on the images of the same size, i.e. 256 × 192 pixel. Therefore, due to the fastest denoising time, the proposed method is suitable for all the applications in which the high number of scans needs to be processed simultaneously.

IV. CONCLUSION AND FUTURE WORK

In this paper, new deep learning schemes for OCT image denoising based on U-Net architecture are proposed. The novel idea consists in specific preprocessing of the input dataset in order to generate pairs of clean and noisy images for

the training of the neural network models. The proposed solutions were evaluated taking into account quantitative metrics concerning image quality as well as the impact of applying the presented denoising scheme in the eye disease classification. The conducted experiments prove that the new methodology can be useful in various OCT image analysis tasks. An important feature of this method is that the processing time was decreased significantly. In our study, we achieved a solution that is 90 times faster in comparison to standard BM3D algorithm while at the same time provides enhanced final image quality. We used a comprehensive dataset of 21926 OCT scans covering both healthy and pathological cases that were collected from 869 patients (1639 eyes) for the creation of denoising models. In turn, the dataset used for disease classification consists of 25697 scans collected from 1910 patients (2953 eyes) that are independent of the previous dataset.

The main limitation of the presented approach consists in the need for making additional preprocessing step (BM3D filtering) before the training of the model might be performed. This might be time-consuming depending on the denoising method that was selected. However, this process is performed only once and the benefits obtained later outweigh these limitations - the processing time of the already trained model is constantly low and satisfying.

Further work will include the application of the proposed methodology in other problems related to OCT image analysis, e.g. in OCT layer segmentation task. Additionally, other referenced image filtering methods and neural network architectures could be tested.

APPENDIX A LIST OF ABBREVIATIONS

See Table 9.

TABLE 9. List of abbreviations.

Abbreviation	Description
AMD	age-related macular degeneration
AUC	area under the ROC curve
BM3D	block-matching and 3D filtering
CNN	convolutional neural network
DR	Diabetic Retinopathy
ERM	epiretinal membrane
fvPED	fibrovascular pigment epithelial detachment
HRF	hyperreflective foci
MH	macular hole
MpH	macular pseudo hole
MSE	mean squared error
MS-SSIM	multi-scale structural similarity
OCT	optical coherence tomography
ONH	optic nerve head
PSNR	peak signal to noise ratio
RD	retinal detachment
RNFL	retinal nerve fiber layer
ROC	receiver operating characteristic
SSIM	structural similarity
sPED	serous pigment epithelial detachment

REFERENCES

- [1] W. Drexler, U. Morgner, R. Ghanta, F. Kärtner, J. Schuman, and A. Fujimoto, "Ultrahigh-resolution ophthalmic optical coherence tomography," *Nature Med.*, vol. 7, pp. 502–507, May 2001.
- [2] M. Adhi and J. Duker, "Optical coherence tomography—Current and future applications," *Current Opinion Ophthalmol.*, vol. 24, pp. 213–221, May 2013.
- [3] R. Maldonado, P. Mettu, M. El-Dairi, and M. Bhatti, "The application of optical coherence tomography in neurologic diseases," *Neurol., Clin. Pract.*, vol. 5, pp. 460–469, Oct. 2015.
- [4] J. Villard, K. Cheruku, and M. Feldman, "Applications of optical coherence tomography in cardiovascular medicine, Part I," *J. Nucl. Cardiol.*, vol. 16, pp. 287–303, Mar. 2009.
- [5] L. Vignali, E. Solinas, and E. Emanuele, "Research and clinical applications of optical coherence tomography in invasive cardiology: A review," *Current Cardiol. Rev.*, vol. 10, no. 4, pp. 369–376, Jul. 2014.
- [6] B. Wan, C. Ganier, X. Du-Harpur, N. Harun, F. Watt, R. Patalay, and A. Lynch, "Applications and future directions for optical coherence tomography in dermatology," *Brit. J. Dermatol.*, vol. 184, pp. 1–10, Sep. 2020.
- [7] J. G. Fujimoto, "Optical coherence tomography for ultrahigh resolution in vivo imaging," *Nature Biotechnol.*, vol. 21, no. 11, pp. 1361–1367, Nov. 2003.
- [8] P. H. Tomlins and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *J. Phys. D, Appl. Phys.*, vol. 38, no. 15, pp. 2519–2535, 2005.
- [9] D. Marks, T. Ralston, and S. Boppart, *Data Analysis and Signal Post-processing for Optical Coherence Tomography*. Berlin, Germany: Springer, Jan. 2008.
- [10] F. J. de Boer, *Spectral/Fourier Domain Optical Coherence Tomography*. Cham, Switzerland: Springer, 2015, pp. 165–193.
- [11] O. Liba, M. D. Lew, E. D. SoRelle, R. Dutta, D. Sen, D. M. Moshfeghi, S. Chu, and A. de la Zorda, "Speckle-modulating optical coherence tomography in living mice and humans," *Nature Commun.*, vol. 8, no. 1, p. 15845, Jun. 2017.
- [12] Y. Du, G. Liu, G. Feng, and Z. Chen, "Speckle reduction in optical coherence tomography images based on wave atoms," *J. Biomed. Opt.*, vol. 19, no. 5, May 2014, Art. no. 056009.
- [13] J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.*, vol. 4, no. 1, pp. 95–105, 1999.
- [14] M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," *Opt. Lett.*, vol. 25, pp. 545–547, May 2000.
- [15] S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, G. Subramanian, L. Zhang, S. Perera, J.-M. Mari, K. S. Chin, T. A. Tun, N. G. Strouthidis, T. Aung, A. H. Thiéry, and M. J. A. Girard, "DRUNET: A dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images," *Biomed. Opt. Exp.*, vol. 9, no. 7, p. 3244, 2018.
- [16] D. Xiang, H. Tian, X. Yang, F. Shi, W. Zhu, H. Chen, and X. Chen, "Automatic segmentation of retinal layer in OCT images with choroidal neovascularization," *IEEE Trans. Image Process.*, vol. 27, no. 12, pp. 5880–5891, Dec. 2018.
- [17] S. Asrani, L. Essaid, B. D. Alder, and C. Santiago-Turla, "Artifacts in spectral-domain optical coherence tomography measurements in glaucoma," *JAMA Ophthalmol.*, vol. 132, no. 4, p. 396, Apr. 2014.
- [18] M. Balasubramanian, C. Bowd, G. Vizzeri, R. Weinreb, and L. Zangwill, "Effect of image quality on tissue thickness measurements obtained with spectral domain-optical coherence tomography," *Opt. Exp.*, vol. 17, pp. 4019–4036, Apr. 2009.
- [19] T. Vajaranant, R. Anderson, R. Zelkha, C. Zhang, J. Wilensky, D. Edward, and M. Shahidi, "The relationship between macular cell layer thickness and visual function in different stages of glaucoma," *Eye*, vol. 25, pp. 612–618, Feb. 2011.
- [20] M. Pircher, E. Gotzinger, R. Leitgeb, A. Fercher, and C. Hitzenberger, "Speckle reduction in optical coherence tomography by frequency compounding," *J. Biomed. Opt.*, vol. 8, pp. 565–569, Aug. 2003.
- [21] J. Schmitt, "Restoration of optical coherence images of living tissue using the clean algorithm," *J. Biomed. Opt.*, vol. 3, pp. 66–75, Jan. 1998.
- [22] T. M. Jørgensen, J. Thomadsen, U. Christensen, W. Soliman, and B. Sander, "Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration—Method and clinical examples," *J. Biomed. Opt.*, vol. 12, no. 4, 2007, Art. no. 041208.

- [23] D. Alonso-Caneiro, S. A. Read, and M. J. Collins, "Speckle reduction in optical coherence tomography imaging by affine-motion image registration," *J. Biomed. Opt.*, vol. 16, no. 11, 2011, Art. no. 116027.
- [24] N. Itfimia, E. B. Bouma, and J. G. Tearney, "Speckle reduction in optical coherence tomography by 'path length encoded' angular compounding," *J. Biomed. Opt.*, vol. 8, no. 2, pp. 260–263, 2003.
- [25] B. Kennedy, T. Hillman, A. Curatolo, and D. Sampson, "Speckle reduction in optical coherence tomography by strain compounding," *Opt. Lett.*, vol. 35, pp. 2445–2447, Jul. 2010.
- [26] W. Wu, O. Tan, R. R. Pappuru, H. Duan, and D. Huang, "Assessment of frame-averaging algorithms in OCT image analysis," *Ophthalmic Surg. Lasers Imag. Retina*, vol. 44, no. 2, pp. 75–168, 2013.
- [27] Y. Liu, H. Simavli, C. J. Que, J. L. Rizzo, E. Tsikata, R. Maurer, and T. C. Chen, "Patient characteristics associated with artifacts in spectralis optical coherence tomography imaging of the retinal nerve fiber layer in glaucoma," *Amer. J. Ophthalmol.*, vol. 159, no. 3, pp. 565–576, Mar. 2015.
- [28] A. Baghaie, R. D'Souza, and Z. Yu, "State-of-the-art in retinal optical coherence tomography image analysis," *Quant. Imag. Med. Surg.*, vol. 5, pp. 603–617, Aug. 2015.
- [29] M. Szkulmowski, I. Gorczynska, D. Szlag, M. Sylwestrzak, A. Kowalczyk, and M. Wojtkowski, "Efficient reduction of speckle noise in optical coherence tomography," *Opt. Exp.*, vol. 20, pp. 1337–1359, Jan. 2012.
- [30] M. Esmacili, A. Dehnavi, H. Rabbani, and F. Hajizadeh, "Speckle noise reduction in optical coherence tomography using two-dimensional curvelet-based dictionary learning," *J. Med. Signals Sensors*, vol. 7, pp. 86–91, Apr. 2017.
- [31] Z. Jian, Z. Yu, L. Yu, B. Rao, Z. Chen, and B. Tromberg, "Speckle attenuation in optical coherence tomography by curvelet shrinkage," *Opt. Lett.*, vol. 34, pp. 1516–1518, Jun. 2009.
- [32] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.
- [33] J. Cheng, D. Tao, Y. Quan, D. W. K. Wong, G. C. M. Cheung, M. Akiba, and J. Liu, "Speckle reduction in 3D optical coherence tomography of retina by A-scan reconstruction," *IEEE Trans. Med. Imag.*, vol. 35, no. 10, pp. 2270–2279, Oct. 2016.
- [34] L. Fang, S. Li, R. P. McNabb, Q. Nie, A. N. Kuo, C. A. Toth, J. A. Izatt, and S. Farsiu, "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," *IEEE Trans. Med. Imag.*, vol. 32, no. 11, pp. 2034–2049, Nov. 2013.
- [35] R. Kafieh, H. Rabbani, and I. Selesnick, "Three dimensional data-driven multi scale atomic representation of optical coherence tomography," *IEEE Trans. Med. Imag.*, vol. 34, no. 5, pp. 1042–1062, May 2015.
- [36] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3450–3459, Dec. 2012.
- [37] J. Aum, J.-H. Kim, and J. Jeong, "Effective speckle noise suppression in optical coherence tomography images using nonlocal means denoising filter with double Gaussian anisotropic kernels," *Appl. Opt.*, vol. 54, no. 13, p. D43, 2015.
- [38] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [39] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Jan. 2013.
- [40] D. Thapa, K. Raahemifar, and V. Lakshminarayanan, "Reduction of speckle noise from optical coherence tomography images using multi-frame weighted nuclear norm minimization method," *J. Mod. Opt.*, vol. 62, pp. 1–9, May 2015.
- [41] B. Chong and Y.-K. Zhu, "Speckle reduction in optical coherence tomography images of human finger skin by wavelet modified BM3D filter," *Opt. Commun.*, vol. 291, pp. 461–469, Mar. 2013.
- [42] S. Chitchian, M. A. Mayer, A. R. Boretzky, F. J. van Kuijk, and M. Motamedi, "Retinal optical coherence tomography image enhancement via shrinkage denoising using double-density dual-tree complex wavelet transform," *J. Biomed. Opt.*, vol. 17, no. 11, Nov. 2012, Art. no. 116009.
- [43] A. Ozcan, A. Bilencu, E. A. Desjardins, E. B. Bouma, and J. G. Tearney, "Speckle reduction in optical coherence tomography images using digital filtering," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 7, pp. 1901–1910, 2007.
- [44] R. Bernardes, C. Maduro, P. Serranho, A. Araújo, S. Barbeiro, and J. Cunha-Vaz, "Improved adaptive complex diffusion despeckling filter," *Opt. Exp.*, vol. 18, pp. 24048–24059, Nov. 2010.
- [45] A. Wong, A. Mishra, K. Bizheva, and D. Clausi, "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery," *Opt. Exp.*, vol. 18, pp. 8338–8352, Apr. 2010.
- [46] Y. Hu, J. Ren, J. Yang, R. Bai, and J. Liu, "Noise reduction by adaptive-SIN filtering for retinal OCT images," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, Sep. 2021.
- [47] T. Schlegl, S. M. Waldstein, H. Bogunovic, F. Endstraßer, A. Sadeghipour, A.-M. Philip, D. Podkowinski, B. S. Gerendas, G. Langs, and U. Schmidt-Erfurth, "Fully automated detection and quantification of macular fluid in OCT using deep learning," *Ophthalmology*, vol. 125, no. 4, pp. 549–558, Apr. 2018.
- [48] X. Sui, Y. Zheng, B. Wei, H. Bi, J. Wu, X. Pan, Y. Yin, and S. Zhang, "Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks," *Neurocomputing*, vol. 237, pp. 332–341, May 2017.
- [49] F. G. Venhuizen, B. van Ginneken, B. Liefers, M. J. J. P. van Grinsven, S. Fauser, C. Hoyng, T. Theelen, and C. I. Sánchez, "Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks," *Biomed. Opt. Exp.*, vol. 8, no. 7, p. 3292, 2017.
- [50] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Exp.*, vol. 8, no. 5, p. 2732, 2017.
- [51] D. Lu, M. Heisler, S. Lee, G. W. Ding, E. Navajas, M. V. Sarunic, and M. F. Beg, "Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network," *Med. Image Anal.*, vol. 54, pp. 100–110, May 2019.
- [52] A. G. Roy, S. Conjeti, S. P. K. Karri, D. Sheet, A. Katouzian, C. Wachinger, and N. Navab, "ReLayNet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Exp.*, vol. 8, no. 8, p. 3627, 2017.
- [53] S. K. Devalla, K. S. Chin, J.-M. Mari, T. A. Tun, N. G. Strouthidis, T. Aung, A. H. Thiéry, and M. J. A. Girard, "A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head," *Investigative Ophthalmol. Vis. Sci.*, vol. 59, no. 1, p. 63, Jan. 2018.
- [54] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," *Biomed. Opt. Exp.*, vol. 8, no. 2, p. 579, 2017.
- [55] M. Awais, H. Müller, T. B. Tang, and F. Meriaudeau, "Classification of SD-OCT images using a deep learning approach," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2017, pp. 489–492.
- [56] P. Prahns, V. Radeck, C. Mayer, Y. Cvetkov, N. Cvetkova, H. Helbig, and D. Maerker, "OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications," *Graefes Arch. Clin. Experim. Ophthalmol.*, vol. 256, pp. 1–8, Jan. 2018.
- [57] C. S. Lee, D. M. Baughman, and A. Y. Lee, "Deep learning is effective for classifying normal versus age-related macular degeneration OCT images," *Ophthalmol. Retina*, vol. 1, no. 4, pp. 322–327, Jul. 2017.
- [58] S. K. Devalla, G. Subramanian, T. H. Pham, X. Wang, S. Perera, T. A. Tun, T. Aung, L. Schmetterer, A. H. Thiéry, and M. J. A. Girard, "A deep learning approach to denoise optical coherence tomography images of the optic nerve head," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, Oct. 2019.
- [59] B. Qiu, Z. Huang, X. Liu, X. Meng, Y. You, G. Liu, K. Yang, A. Maier, Q. Ren, and Y. Lu, "Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function," *Biomed. Opt. Exp.*, vol. 11, no. 2, pp. 817–830, 2020.
- [60] Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," *Biomed. Opt. Exp.*, vol. 9, no. 11, p. 5129, 2018.
- [61] K. J. Halupka, B. J. Antony, M. H. Lee, K. A. Lucy, R. S. Rai, H. Ishikawa, G. Wollstein, J. S. Schuman, and R. Garnavi, "Retinal optical coherence tomography image enhancement via deep learning," *Biomed. Opt. Exp.*, vol. 9, no. 12, p. 6205, 2018.
- [62] F. Shi, N. Cai, Y. Gu, D. Hu, Y. Ma, Y. Chen, and X. Chen, "DeSpecNet: A CNN-based method for speckle reduction in retinal optical coherence tomography images," *Phys. Med. Biol.*, vol. 64, no. 17, Sep. 2019, Art. no. 175010.

- [63] J. J. Rico-Jimenez, D. Hu, E. M. Tang, I. Oguz, and Y. K. Tao, "Real-time OCT image denoising using a self-fusion neural network," *Biomed. Opt. Exp.*, vol. 13, no. 3, p. 1398, 2022.
- [64] Z. Mao, A. Miki, S. Mei, Y. Dong, K. Maruyama, R. Kawasaki, S. Usui, K. Matsushita, K. Nishida, and K. Chan, "Deep learning based noise reduction method for automatic 3D segmentation of the anterior of lamina cribrosa in optical coherence tomography volumetric scans," *Biomed. Opt. Exp.*, vol. 10, no. 11, p. 5832, 2019.
- [65] Y. Huang, N. Zhang, and Q. Hao, "Real-time noise reduction based on ground truth free deep learning for optical coherence tomography," *Biomed. Opt. Exp.*, vol. 12, no. 4, p. 2027, 2021.
- [66] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [67] M. Li, R. Idoughi, B. Choudhury, and W. Heidrich, "Statistical model for OCT image denoising," *Biomed. Opt. Exp.*, vol. 8, no. 9, pp. 3903–3917, 2017.
- [68] Y. Daiqiang, Y. Gu, and P. Xue, "Speckle-constrained variational methods for image restoration in optical coherence tomography," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 30, pp. 878–885, May 2013.
- [69] W. Feng, H. Lei, and Y. Gao, "Speckle reduction via higher order total variation approach," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1831–1843, Apr. 2014.
- [70] D. Varadarajan, C. Magnain, M. Fogarty, D. A. Boas, B. Fischl, and H. Wang, "A novel algorithm for multiplicative speckle noise reduction in *ex vivo* human brain OCT images," *NeuroImage*, vol. 257, p. 119304, 2022, doi: 10.1016/j.neuroimage.2022.119304.
- [71] D. Yang and J. Sun, "BM3D-Net: A convolutional neural network for transform-domain collaborative filtering," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 55–59, Jan. 2018.
- [72] *INDOK Project*. Accessed: Jun. 3, 2023. [Online]. Available: <https://www.consultronix.pl/dotacje-unijne/indok>
- [73] J. Immerkær, "Fast noise variance estimation," *Comput. Vis. Image Understand.*, vol. 64, no. 2, pp. 300–302, Sep. 1996.
- [74] *Scikit-Learn Package*. Accessed: Jun. 3, 2023. [Online]. Available: <https://scikit-learn.org/>
- [75] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351, 2015, pp. 234–241.
- [76] *BM3D Python Package*. Accessed: Jun. 3, 2023. [Online]. Available: <https://pypi.org/project/bm3d/>
- [77] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [78] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, 2003, pp. 1398–1402.
- [79] X.-J. Mao, C. Shen, and Y. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. NIPS*, 2016, pp. 1–16.
- [80] *Keras Architectures in Tensorflow*. Accessed: Jun. 3, 2023. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/applications



PIOTR BOGACKI was born in Kraków, Poland, in 1989. He received the B.E. and M.S. degrees in telecommunications engineering from the AGH University of Science and Technology, Kraków, in 2012 and 2014, respectively, where he is currently pursuing the Industrial Ph.D. degree in computer science and telecommunications.

He has participated in the Erasmus exchange program, within which he spent the last academic year with Universitat Politècnica de Valencia, Valencia, Spain, in 2012 and 2013. In 2014, he realized a professional internship with Capgemini company, Valencia, where he developed database and business intelligence systems. From 2014 to 2018, he was a Software Engineer with Samsung R&D Poland, where he was responsible for the development of 4G and 5G systems. Since 2018, he has been an Expert Algorithm Development Engineer in automotive sector with Aptiv Poland, Kraków, where he has been responsible for the development of active safety vision features. Since 2011, he has been involved in seven research projects. His research interests include image processing and video analysis.



ANDRZEJ DZIECH received the M.Sc. and Ph.D. degrees from the Institute of Electrical Engineering, Saint Petersburg, in 1970 and 1973, respectively, and the D.Sc. degree from the Technical University of Poznan, in 1978.

He is currently a Full Professor with the Institute of Telecommunications, AGH University of Science and Technology, Kraków, Poland. He was a Visiting Professor with Imperial College London, U.K., from 1982 to 1984, and University of Wuppertal, Germany. He is the author of six books, and nearly 230 publications. He was a supervisor of 19 Ph.D. students. His research interests include digital communications, image and data processing, data compression, information and coding theory, random signals, computer communications networks, and signal processing.

Prof. Dziech has been honored four times by the Ministry of Science and Education of Poland for his research achievements. In 2018, he was awarded by the Prime Minister of Poland for his outstanding research results. He actively participated in numerous international research projects, e.g., INDECT, SCISSOR, and ECHO.

...