Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Learnable despeckling framework for optical coherence tomography images

Saba Adabi
Elaheh Rashedi
Anne Clayton
Hamed Mohebbi-Kalkhoran
Xue-wen Chen
Silvia Conforto
Mohammadreza Nasiriavanaki

SPIE.

Saba Adabi, Elaheh Rashedi, Anne Clayton, Hamed Mohebbi-Kalkhoran, Xue-wen Chen, Silvia Conforto, Mohammadreza Nasiriavanaki, "Learnable despeckling framework for optical coherence tomography images," *J. Biomed. Opt.* **23**(1), 016013 (2018), doi: 10.1117/1.JBO.23.1.016013.

Learnable despeckling framework for optical coherence tomography images

Saba Adabi,[a,b] Elaheh Rashedi,[c] Anne Clayton,[a] Hamed Mohebbi-Kalkhoran,[a] Xue-wen Chen,[c] Silvia Conforto,[b] and Mohammadreza Nasiriavanaki[a,d,e,*]
[a]Wayne State University, Department of Biomedical Engineering, Detroit, Michigan, United States
[b]Roma Tre University, Department of Applied Electronics, Rome, Italy
[c]Wayne State University, Department of Computer Science, Detroit, Michigan, United States
[d]Wayne State University, Department of Neurology, Detroit, Michigan, United States
[e]Barbara Ann Karmanos Cancer Institute, Detroit, Michigan, United States

Abstract. Optical coherence tomography (OCT) is a prevalent, interferometric, high-resolution imaging method with broad biomedical applications. Nonetheless, OCT images suffer from an artifact called speckle, which degrades the image quality. Digital filters offer an opportunity for image improvement in clinical OCT devices, where hardware modification to enhance images is expensive. To reduce speckle, a wide variety of digital filters have been proposed; selecting the most appropriate filter for an OCT image/image set is a challenging decision, especially in dermatology applications of OCT where a different variety of tissues are imaged. To tackle this challenge, we propose an expandable learnable despeckling framework, we call LDF. LDF decides which speckle reduction algorithm is most effective on a given image by learning a figure of merit (FOM) as a single quantitative image assessment measure. LDF is learnable, which means when implemented on an OCT machine, each given image/image set is retrained and its performance is improved. Also, LDF is expandable, meaning that any despeckling algorithm can easily be added to it. The architecture of LDF includes two main parts: (i) an autoencoder neural network and (ii) filter classifier. The autoencoder learns the FOM based on several quality assessment measures obtained from the OCT image including signal-to-noise ratio, contrast-to-noise ratio, equivalent number of looks, edge preservation index, and mean structural similarity index. Subsequently, the filter classifier identifies the most efficient filter from the following categories: (a) sliding window filters including median, mean, and symmetric nearest neighborhood, (b) adaptive statistical-based filters including Wiener, homomorphic Lee, and Kuwahara, and (c) edge preserved patch or pixel correlation-based filters including nonlocal mean, total variation, and block matching three-dimensional filtering. *© 2018 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.23.1.016013]

Keywords: optical coherence tomography; speckle; image denoising; image enhancement.

Paper 170341R-APPEAL received May 25, 2017; accepted for publication Jan. 4, 2018; published online Jan. 24, 2018.

1 Introduction

Optical coherence tomography (OCT) is a noninvasive, low coherent interferometry technique that uses backscattered light to generate tomographic images of tissue microstructures.[1] With advances in imaging technology and optical devices, OCT currently has extensive biomedical applications in various fields including dermatology.[1,2] Although OCT offers high-resolution, three-dimensional (3-D) images of tissues with 2- to 15-μm resolution and 1- to 2-mm imaging depth,[3] OCT images suffer from a grainy texture artifact, speckle, due to the broadband/low coherent light source used in the configuration of OCT.[3] Speckle is formed, if the out-of-phase backscattered signals reach the detector within the coherence time of the laser, and it is the summation of multiple optical backscattered wavefields from the sample.[4] Speckle reduces the image quality, e.g., spatial resolution of the borders and edges in the image.[5] Since speckle is an artifact that carries submicron structural information, devising an appropriate speckle reduction algorithm becomes a challenging task.[6] Speckle reduction methods can be categorized into hardware-based methods[7] and software-based methods.[8,9] Compounding techniques are the most common hardware-based methods. In these methods, the imaging is performed several times and each time one system parameter is altered, e.g., angle of light illumination (angular compounding), polarization of the incident light (polarization compounding), and central wavelength of the light source (spectral compounding).[4,7,9] Software-based methods are digital filters relying on a mathematical model of the speckle and are implemented based upon the local or overall statistics of the image. Among speckle reduction digital filtering methods, the most common ones, such as averaging and median, are time efficient; however, they reduce the spatial resolution in the image. Improved performance is provided by adaptive digital filters including Lee,[10] Kuwahara,[11] and Wiener filter.[12] Moreover, wavelet-based filtering methods[9] and a diffusion-based filter with fuzzy logic thresholding as transform domain techniques have demonstrated adequate results.[13] Some of these techniques are compared in Ref. 8 and authors concluded that the combination of enhanced Lee method and Wiener filter could significantly improve the quality of OCT images of *ex vivo* bovine retina. The other filters were developed based on an A-scan reconstruction procedure[14]

*Address all correspondence to: Mohammadreza Nasiriavanaki, E-mail: mrn.avanaki@wayne.edu

1083-3668/2018/$25.00 © 2018 SPIE

or Bayesian estimation[15] for OCT image speckle reduction. Several other studies have been carried out on speckle reduction by tuning the internal parameters in a particular despeckling algorithm.[16] Relying on more advanced filtering methods, several speckle reduction methods were developed based on the total variation (TV) concept,[17] nonlocal mean (NLM) filtering,[17] and block-matching 3-D filtering (BM3D).

A tissue's cellular specification can be represented by optical and textural features extracted from OCT images. Together with the attenuation coefficients,[18] textural features can provide valuable information about the texture and scattering properties of a biological tissue.[19,20] We introduced an expandable despeckling framework, named learnable despeckling framework (LDF). LDF decides which speckle reduction algorithm is most effective for a given image by learning a single quantitative image assessment measure: the figure of merit (FOM). Among several algorithms developed for speckle reduction in OCT images, we utilized the above-mentioned digital filters to evaluate our proposed framework. With any given image to LDF, it is retrained and its performance improved. Any other algorithm can easily be added to this framework; this is the expandable characteristic of LDF. LDF's architecture is composed of two main parts; an autoencoder neural network and a filter classifier. In our proposed scheme, initially the autoencoder learns to construct a weighted FOM measure based on quality assessment measures, i.e., signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR),[8,21–23] equivalent number of looks (ENL), mean structural similarity (MSSIM) index, and edge preservation index (EPI), extracted from filtered OCT images in the following categories: (a) sliding window filters including median, mean, and SNN; (b) adaptive statistical-based filters including Wiener, homomorphic Lee, and Kuwahara; and (c) edge preserved patch or pixel correlation-based filters including NLM, TV, and BM3D. Then, the filter classifier identifies the most efficient filter.

2 Materials and Methods

In Sec. 2.1, we present the digital filtering methods used in this study. In Sec. 2.2, we describe statistical features that can be extracted from an OCT image. In Sec. 2.3, we explain the image quality assessment measures that are utilized to evaluate the quality of the filtered images. Following that, in Sec. 2.4, we introduce the architecture of the LDF in details. The OCT system specifications are given in Sec. 2.5.

2.1 *Digital Filtering Methods*

Digital filters are means to implement mathematical operations characterized by its transfer functions or, equivalently, on a sampled signal or image, to improve its quality. A speckle reduction method is designed and implemented based on the statistical characteristics of speckle. There are three main classes of digital filters such as sliding window, adaptive statistical based, and edge preserved patch or pixel correlation-based filters. We explored 25 filters. Among them, filters #1 to #4 are median filter with window sizes 3, 5, 7, and 9 pixels, respectively; filters #5 to #8 are averaging filter with window sizes 3, 5, 7, and 9 pixels, respectively; filters #9 to #12 are symmetric nearest neighborhood (SNN)[16] with window sizes 3, 5, 7, and 9 pixels, respectively; filters #13 and #14 are Kuwahara[24] with window sizes of 3 and 5 pixels, respectively; filters #15 to #18 are adaptive Wiener filter with window sizes 3, 5, 7, and 9 pixels, respectively;[25] filters #19 to #22 are Lee filter with window sizes 3, 5, 7, and 9 pixels, respectively; filter #23 is a pixel-wise NLM filter;[26] filter #24 is a TV[16] filter; and filter #25 is a BM3D.[27]

2.1.1 *Sliding window filters*

This class of filter includes mean, median, and SNN. These filters are time-efficient and can be used in real-time speckle reduction applications, such as video-rate OCT imaging. Although they effectively reduce speckle noise in the OCT image, they smooth edges in the image and create blurriness. Mean filter is a linear convolutional low-pass filter. In this filter, a pixel value is replaced by the average of its neighboring pixel values. In median filtering a pixel value in a window, $M \times N$ pixels, is replaced by the value of the middle pixel in a vector of pixel values sorted in an ascending order.[28] This nonlinear filter is more robust than the mean filter and preserves edges more effectively. SNN is considered as an edge preserving sliding window speckle reduction method. In SNN, initially the opposite pairs of pixels in the support are compared and replaced with the pixel value that is closest to the central pixel value.[16] Each pixel value is then replaced by the average of processed pixel values in the support.

2.1.2 *Adaptive statistical-based filters*

This class of despeckling filter includes Kuwahara filter,[11] adaptive Wiener filter, and Lee filter and utilizes statistical features e.g., mean and variance, extracted from the image or a part of the image. Kuwahara works by dividing the support into four subregions.[24] The central pixel is replaced with the average of the quadrants with the lowest variance. Wiener filter tailors itself to image local mean and local variance, i.e., the larger the variance, the less smoothing is applied.[12] Lee filter is an adaptive filter that determines each pixel value according to the weighted sum of the center pixel value; the local statistics (mean and variance) calculated in a square kernel surrounding the pixel with a minimum mean square error (MMSE) approach.[29]

2.1.3 *Patch or pixel correlation-based filters*

This class of despeckling filter, includes NLM, TV, and BM3D, are based on high inter- or intra-correlations among nearby pixels or patch of pixels. The NLM filter algorithm changes the value of the target pixel by taking the average value of all or selected pixels in the image and weighting them based on their similarity to the target pixel. NLM filters are known to preserve the textures.[26] TV filters are based on an edge preserving TV regularization process that relies on the fact that signals with excessive detail have high TV, implying a large integral of absolute gradient for the signal.[30] TV provides a close match to the ground truth image. TV efficiently suppresses the noise while preserving the image details. BM3D is a collaborative filtering method designed considering the fact that the image has a locally sparse representation in the transform domain.[27] The procedure begins with grouping similar image patches into three-dimensional (3-D) groups. Then, a 3-D linear transformation is applied on the image and a shrinkage procedure is performed. Following this process, an inverse transformation is applied on spectrum coefficients. Combining the patches results in an estimation of the ground truth image. At the end, a Wiener filter is used to form the final denoised image.[31]

2.2 *Image Textural and Optical Features*

To quantify tissue properties, 27 features including 26 texture features[32,33] and 1 optical property are extracted from OCT images. For each image, 6 first-order statistical features including mean, variance, skewness, median, kurtosis, and entropy are calculated. Twenty features from the gray-level co-occurrence matrix (GLCM)[34] i.e., homogeneity, contrast, energy, entropy, and correlation in four directions, 0 deg, 45 deg, 90 deg, and 135 deg, are calculated as textures.[35] The optical property calculated for the OCT image is the attenuation coefficient. We used Vermeers' method[36] to calculate the attenuation coefficient for each pixel in the OCT intensity image. We utilized principal components analysis (PCA) algorithm to reduce the dimension of the features (from 27 to 5) with the least loss of information. PCA is a projection-based method that reduces the computational complexity through construction of orthogonal principal components.[37]

2.3 *Image Quality Assessment Measures*

The performance of the filtering methods is assessed using well-established objective assessment measures, including SNR, CNR,[8,21–23] ENL, MSSIM, and EPI.[9,38,39] SNR compares the

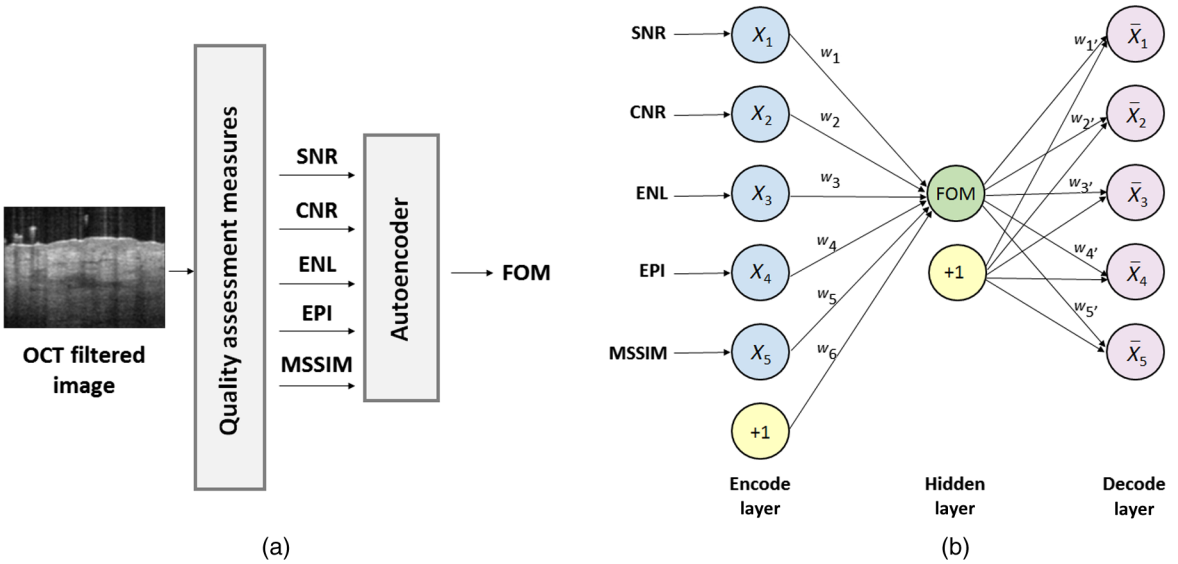

Fig. 1 (a) Block diagram of FOM calculation, (b) the structure of autoencoder, X_1 to X_5 are nodes of the encode layer, w_1 to w_6 and w_1' to w_5' are the weigh parameters, and $\bar{X}_1$ to $\bar{X}_5$ are nodes of the decode layer. FOM is a node of the hidden layer.

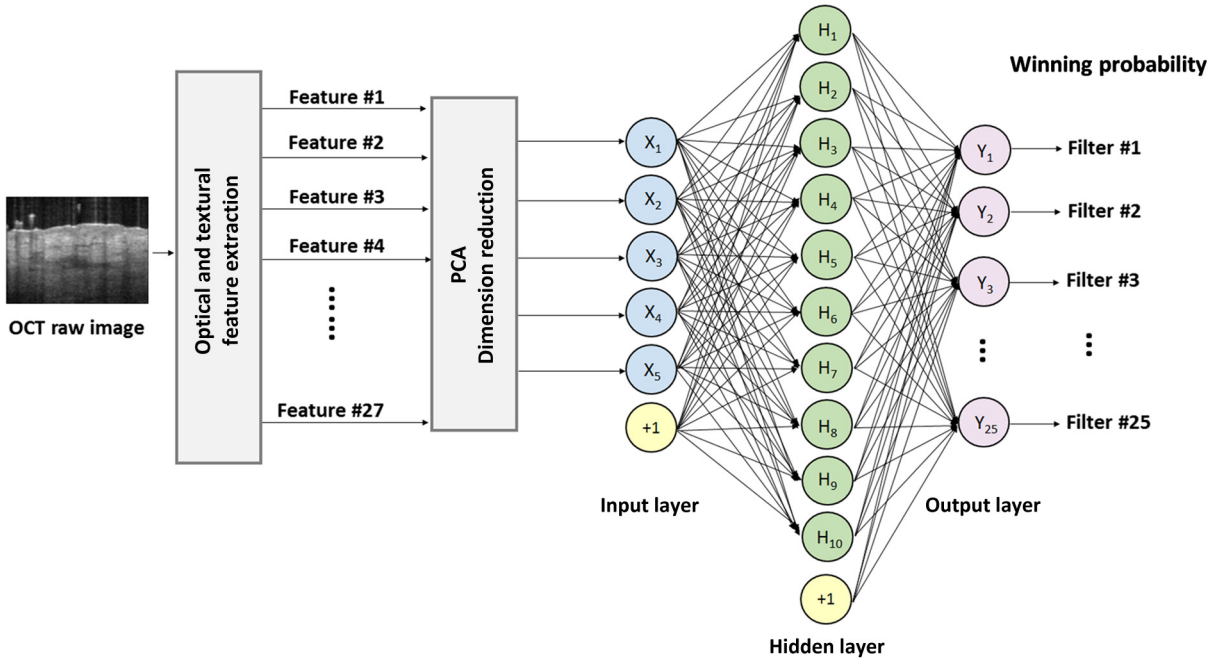

Fig. 2 The architecture of the classifier. The ANN is trained to select the best filter for the given input OCT image. X_1 to X_5 are input nodes, H_1 to H_{10} are hidden nodes, and Y_1 to Y_{25} are output nodes.

signal of an object in the OCT image to its background noise. CNR is a measure of the signal fluctuations to the noise. The definition of SNR and CNR is given in Eqs. (1) and (2), respectively. Both SNR and CNR are calculated over a number of ROIs (10 in this study)

$$\mathrm{SNR} = 10\log_{10}\left[\frac{\mathrm{median}(I_{\mathrm{lin}}^2)}{\sigma_{\mathrm{lin}}^2}\right], \tag{1}$$

$$\mathrm{CNR} = \frac{1}{R}\left[\sum_{r=1}^{R}\frac{(\mu_r - \mu_b)}{\sqrt{\sigma_{r^2} + \sigma_{b^2}}}\right], \tag{2}$$

where I_{lin} is the linear magnitude of pixel intensities in a homogeneous region of interest (ROI) where there is tissue, σ_{lin}^2 is the variance of the pixel values in the background region, μ_b and σ_b represent the mean and variance of the background region, μ_r and σ_r^2 represent the mean and variance of the r'th ROI.[21] ENL is a measure of smoothness in a homogeneous ROI and can be calculated as

$$\mathrm{ENL} = \frac{1}{H}\sum_{h}^{H}\mu_h^2\sigma_h^2, \tag{3}$$

where μ_h^2 and σ_h^2 are the mean and variance of homogeneous ROIs (H). The MSSIM index quantifies image quality referring to its structural similarities and is based on local statistic calculations

$$\mathrm{MSSIM} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{[2\mu_{\acute{I}_{(i,j)}}\mu_{\hat{I}_{(i,j)}} + C_1][2\sigma_{\acute{I}_{(i,j)}}\sigma_{\hat{I}_{(i,j)}} + C_2]}{[\mu_{\acute{I}_{(i,j)}}^2 + \mu_{\hat{I}_{(i,j)}}^2 + C_1][\sigma_{\acute{I}_{(i,j)}}^2 + \sigma_{\hat{I}_{(i,j)}}^2 + C_2]}, \tag{4}$$

where M and N are the sizes of the image in transverse directions, I is the original image, and $\hat{I}$ is the despeckled image. $\acute{I}_{(i,j)}$ and $\hat{I}_{(i,j)}$ are derived by convolving the original and despeckled images with a symmetric Gaussian kernel with window size 11 to calculate their local variance and mean, i.e., $\sigma_{\acute{I}_{(i,j)}}$, $\sigma_{\hat{I}_{(i,j)}}$, $\mu_{\acute{I}_{(i,j)}}$, and $\mu_{\hat{I}_{(i,j)}}$. C_1 and C_2 are the constants: $C_1 = 6.502$ and $C_2 = 58.522$.[39] The ground truth image is generated by averaging

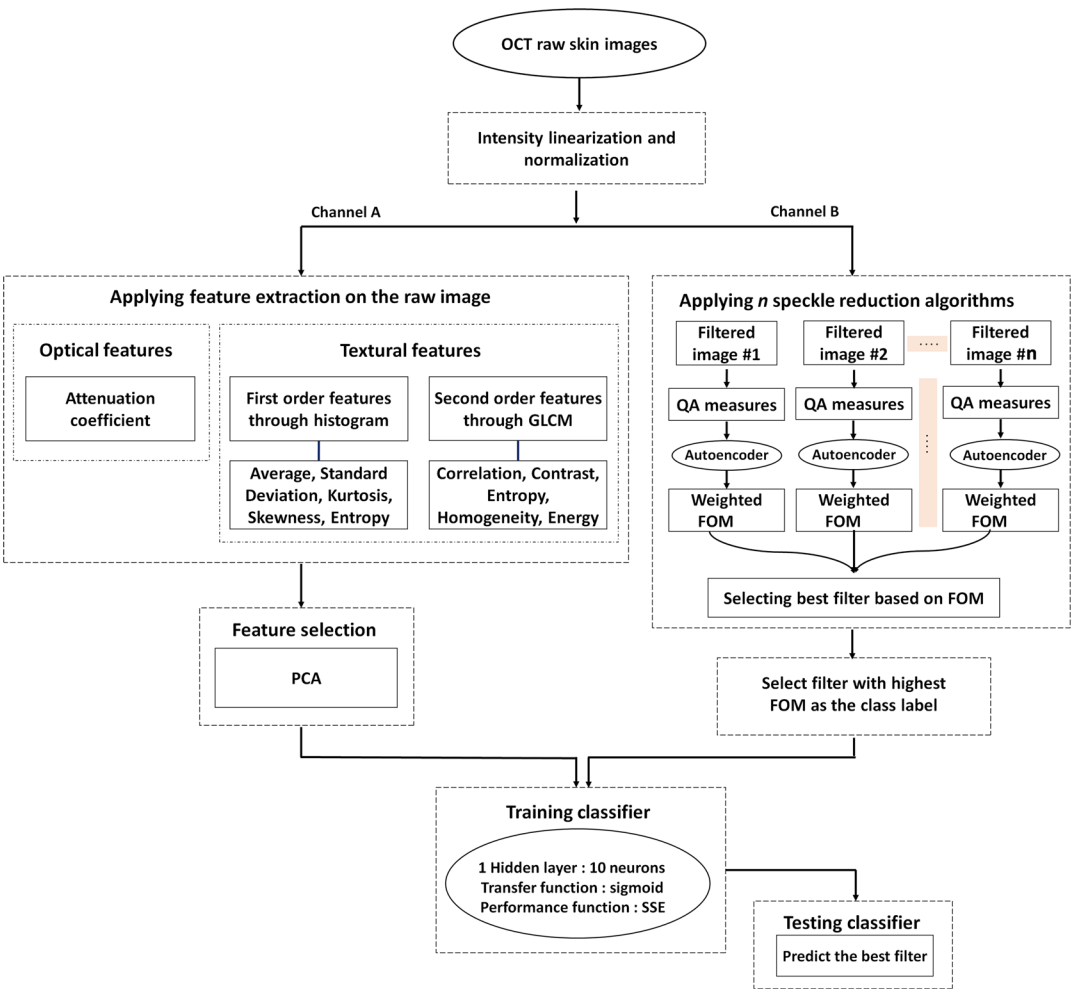

Fig. 3 Block diagram of the LDF algorithm. OCT, optical coherence tomography, QA, quality assessment; FOM, figure of merit; and GLCM, gray-level co-occurrence matrix. QA measures include SNR, CNR, ENL, MSSIM, and EPI, n is the number of filters (in this study $n = 25$), SSE is the sum square error, chosen as the performance function.

170 (can be considered as spatial compounding method[40] since the images are taken from slightly misaligned samples due to imperfect optical scanners used in the OCT machine) B-scan images to calculate MSSIM index. EPI is a correlation-based method that shows how the edges in the image degrade

$$\mathrm{EPI} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[\Delta I_{(i,j)} - \mu_{\Delta I_{(i,j)}}][\Delta \hat{I}_{(i,j)} - \mu_{\hat{I}_{(i,j)}}]}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[\Delta I_{(i,j)} - \mu_{\Delta I_{(i,j)}}][\Delta \hat{I}_{(i,j)} - \mu_{\hat{I}_{(i,j)}}]}}, \quad (5)$$

where I indicates the original image, $\hat{I}$ is the despeckled image, ΔI is an edge detected image with a Laplacian operator, μ and σ are the mean and variance of the image, μ_b and σ_b^2 indicate the mean and variance of the background, respectively.

2.4 *Learnable Despeckling Framework*

The architecture of LDF includes two main parts: (i) an autoencoder neural network and (ii) filter classifier. In Sec. 2.4.1, we explain the architecture of the autoencoder utilized to learn the FOM. In Sec. 2.4.2, we describe the architecture of the classifier that can be trained based on FOM to predict the most effective despeckling filter.

2.4.1 *Autoencoder architecture*

An FOM is defined as a single representative measurement to assess the performance of each filter. In this study, we define the FOM based on a set of five OCT quality measures including SNR, CNR, ENL, EPI, and MSSIM. The goal is to find an FOM to best represent the quality of an image. To do this, we utilized an autoencoder artificial neural network (ANN) with three layers for unsupervised training of FOM. The structure of the autoencoder is shown in Fig. 1. As it is shown, layer 1 contains six neurons, including SNR, CNR, ENL, EPI, MSSIM, and a bias neuron. Layer 2 includes one neuron to estimate the FOM and a biased neuron. Sigmoid transfer function, the stochastic gradient descent optimizer, and MSE as the loss function are used in training the autoencoder. Autoencoders work well (accuracy above 90% and faster convergence) if initially all the weights are the same. Following the work in Ref. 8, for applications similar to ours, the initial weights of one is the best choice.[41] The autoencoder is trained to calculate the FOM by utilizing the quality assessments measures obtained from the filtered images. In this experiment, the final weights of the encoder layer were calculated as: $[w_1, w_2, w_3, w_4, w_5] = [0.1237, 0.2387, 0.0296, 0.1987, 0.4093]$, where $[w_1, w_2, w_3, w_4, w_5]$ corresponds to [SNR, CNR, ENL, EPI, MSSIM]. Based on this experiment, one can conclude that the MSSIM has a more significant effect on the FOM, whereas ENL has less significant effect.

2.4.2 *Classifier architecture*

We used FOM measure to classify the despeckling filters. In this study, we utilized another ANN as the classifier. The classifier predicts the most effective filter (the winner filter) for the given input image. The designed ANN classifier includes three layers, the input layer, the hidden layer, and the output layer. The input layer includes five neurons corresponding to five features extracted from the image (the number of extracted features is initially 27, which is reduced to 5 by utilizing a PCA algorithm). The hidden layer includes 10 neurons. Finally, the output layer includes 25 neurons, which is equal to the number of filters in the experiment. The stochastic gradient descent optimizer and sum square error (SSE) as the loss function is used in training

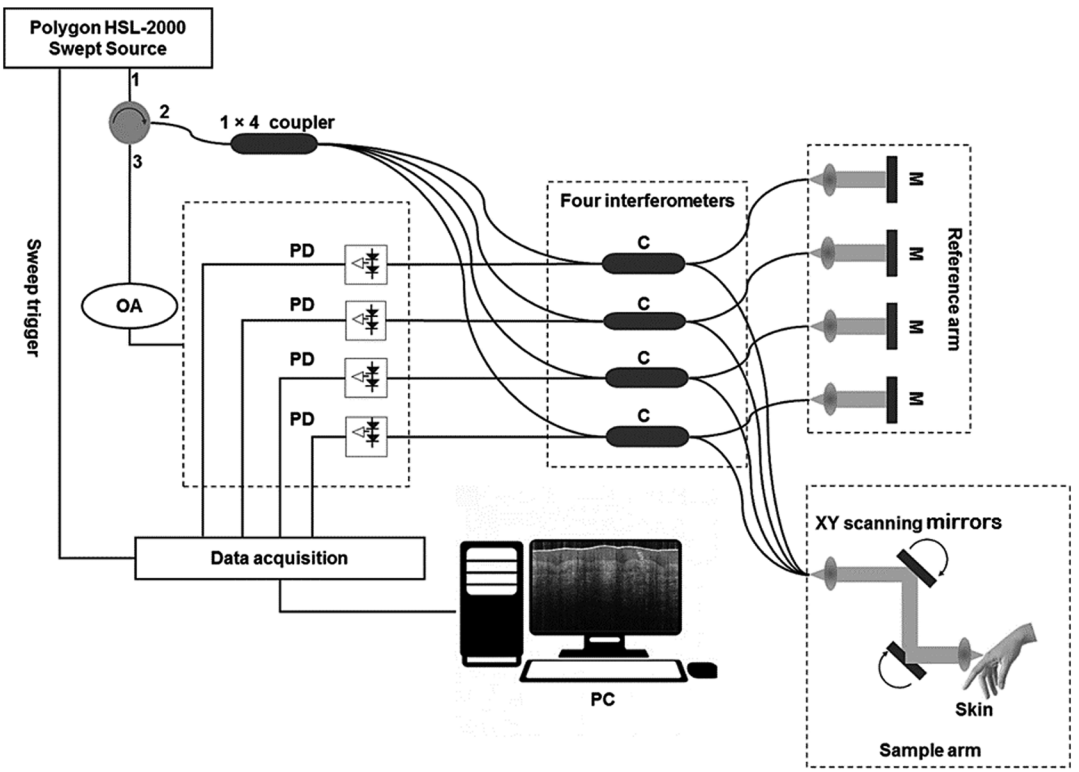

Fig. 4 Schematic diagram of the multibeam swept-source OCT; M, mirror; C, optical coupler; PD, photodetector; and OA, optical attenuator.

the classifier. Figure 2 shows the architecture of the classifier. The value of each neuron in the output layer is a real number between 0 and 1, which represents the probability of the corresponding filter being the winner filter. At the end, the filter with the highest probability will be selected as the winner filter. Before training, the features are normalized, once over each image and once for the ensemble of all images.

The flowchart of the LDF (including classifier's training and testing steps) is given in Fig. 3. We are provided with a set of OCT row skin images. The images are initially linearized and normalized. The processed images are then passed through two parallel channels, A and B. In channel A, we extract the optical and textural features of the images and convert the set of the correlated features to a smaller set of linearly uncorrelated features by utilizing the PCA algorithm. In channel B, we apply all the filtering methods on the image to create a set of filtered images. Then, we calculate the FOM of each filtered image using our pre-trained autoencoder. Thereafter, the filter that achieves the highest FOM is chosen as the winner filter and hence is the class label of the selected image. This process repeats until all images have a class label. At the end, the classifier is trained and tested using the selected features and the class labels.

2.5 *OCT System Specifications*

We use a multibeam swept-source OCT system (SS-OCT) (VivoSight, Michelson Diagnostic™ Inc., United Kingdom) for this study (Fig. 4). The lateral and axial resolutions are 7.5 and 10 μm, respectively. The scan area of the OCT system is 6 mm (width) ×6 mm (length) ×2 mm (depth). A tunable broadband laser source with the central wavelength of 1305 ± 15 nm, successively sweeps through the optical spectrum, leads the light to four separate interferometers, and forms four consecutive confocal gates. The interference spectrum generated by the frequency sweep, over the whole bandwidth in time, is given with respect to frequency. The 10-kHz sweep is the frequency that one reflectivity profile (A-scan) is generated. A B-scan is

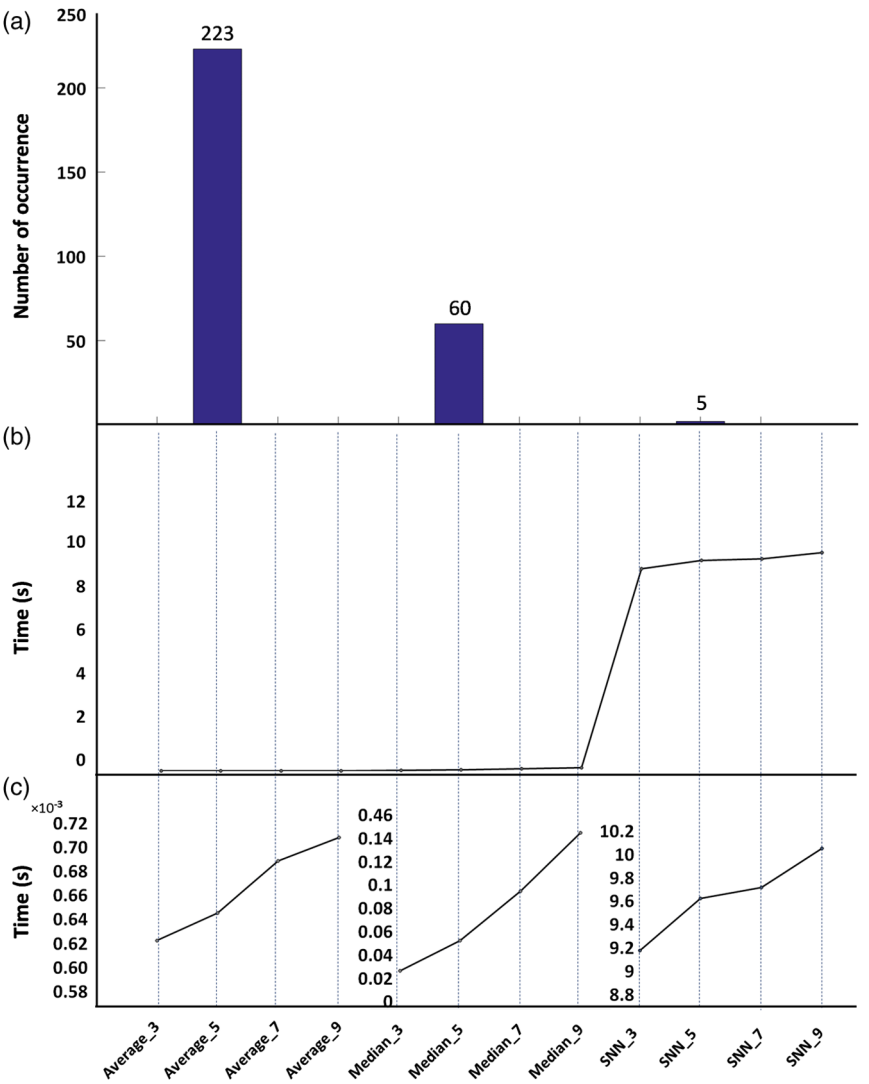

Fig. 5 Histogram of winner filter for 285 OCT images with their corresponding execution time. Sliding window filters are used here. Histogram (a) illustrates the number of images that each filter is winning on, histogram (b) shows running time for each filtering method, and histogram (c) magnifies histogram (b) in three distinct scales to elaborate the running time comparison between different window sizes of average, median, and SNN filters.

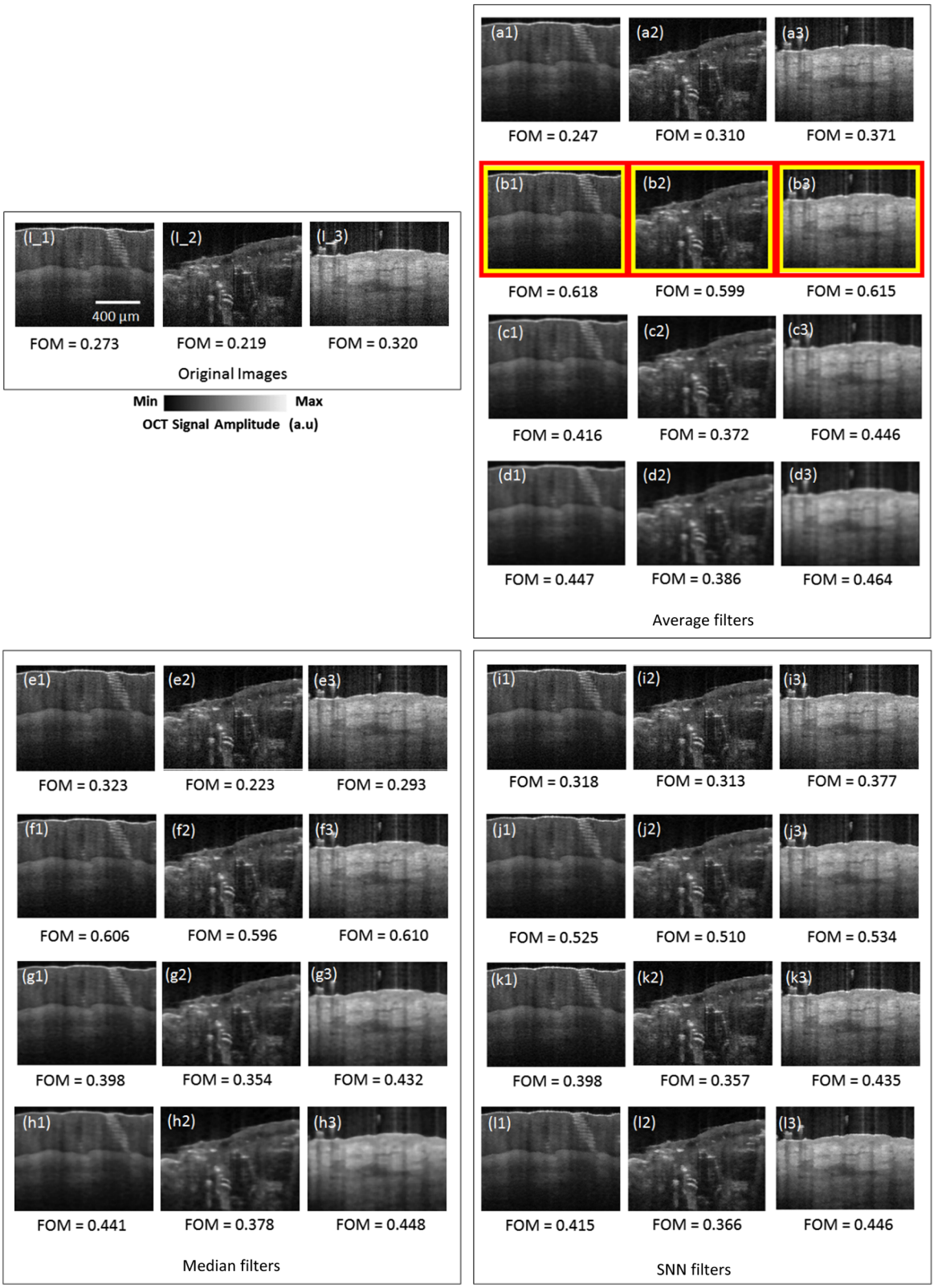

Fig. 6 Results of despeckling using sliding window filters on three test images. Original OCT images taken from (I_1) healthy thumb of a 24-year-old male, (I_2) acne diseased outer arm of a 56-year-old female, (I_3) back of a healthy 25-year-old male, (a–d) despeckled images using averaging with window sizes 3, 5, 7, and 9, respectively, (e–h) despeckled images using median filters with window sizes 3, 5, 7, and 9, respectively, (i–l) despeckled images using SNN filters with window sizes 3, 5, 7, and 9, respectively. The yellow boxes indicate the winner filter based on FOM measure, the red boxes indicate the winner filter chosen by the classifier in sliding window filters subgroup.

then generated by combining several adjacent A-scans for different transversal positions of the incident beam.

Due to the multibeam configuration, our Vivosight OCT can be considered approximately as a discrete dynamic focus OCT,[42] and with a good approximation, these parameters can be neglected. Therefore, compensation for confocal parameter of the lens and for the fall in laser coherence was not performed.

3 Results and Discussion

We design three sets of experiment to demonstrate whether the FOM is a reliable merit and if it could be used as the class label in training the classifier (see Sec. 3.1). The experiments are performed based on three main classes of digital filters, i.e., sliding window, adaptive statistical based, and edge preserved patch or pixel correlation-based filters. Following these experiments, in Sec. 3.2, we use the fivefold cross-validation technique to estimate the accuracy of the trained classifier.

3.1 *Figure of Merit Learning and Validation*

The 27 features are computed from 25 ROIs in each image. For SNR and CNR calculations, 20 ROIs are selected from the tissue region and 10 ROIs from the background region. For computation of other three quality assessment measures, the entire filtered image is used. All the 25 digital filters described in Sec. 2 were implemented in MATLAB® 2016. We used a Dell desktop computer with an Intel Core i7, 3.10 GHz CPU and 8 GB of RAM to implement the algorithms. The OCT machine is an FDA approved system for skin imaging, thus, *in vivo* skin images were collected. Images were acquired from both healthy and diseased individual's skin. OCT images of healthy skin were taken from various body locations, to account for the variety of skin architecture found on the body. Additionally, OCT images of diseased skin were collected, including nonmelanoma skin cancer, psoriasis, and acne. The imaging was performed in Oakwood Clinic, Dearborn, Michigan. The institutional review board at Wayne State University (Independent Investigational Review Board, Detroit, Michigan) approved the study protocol. For training the classifier, the number of input images was $285 \times 25 = 7125$, where 285 is the number of images and 25 is the number of filters. Based on a fivefold cross-validation method, out of 7125 OCT images, $228 \times 25 = 5700$ images were used for training the classifier, and the remaining (i.e., 1425 images) were used for test.[43] The accuracy of the classifier is obtained as 97%. In the following, three experiments are performed based on the three main classes of digital filters, i.e., sliding window, adaptive statistical-based, and edge preserved patch or pixel correlation-based filters.

The histogram of winner filters for 285 test sets in the sliding window filter category as well as their execution time are shown in Fig. 5. According to the graph, the average filter with the window size 5 is the winner filter for despeckling, most of the time. Median filter with the window size 5, and SNN with the window size 5 are the next two winner filters.

In Fig. 6, three original OCT images and their despeckled images using sliding window filters are shown. In the results presented in Fig. 6, the classifier has only learned the sliding window filters subgroup rather than the entire filter pool. The yellow boxes in the figure indicate the winner filters based on FOM criterion. For three test images here, the average filter with the window size 5 was chosen as the winner filter. The red boxes indicate the winner filter chosen by the classifier in sliding window filters subgroup.

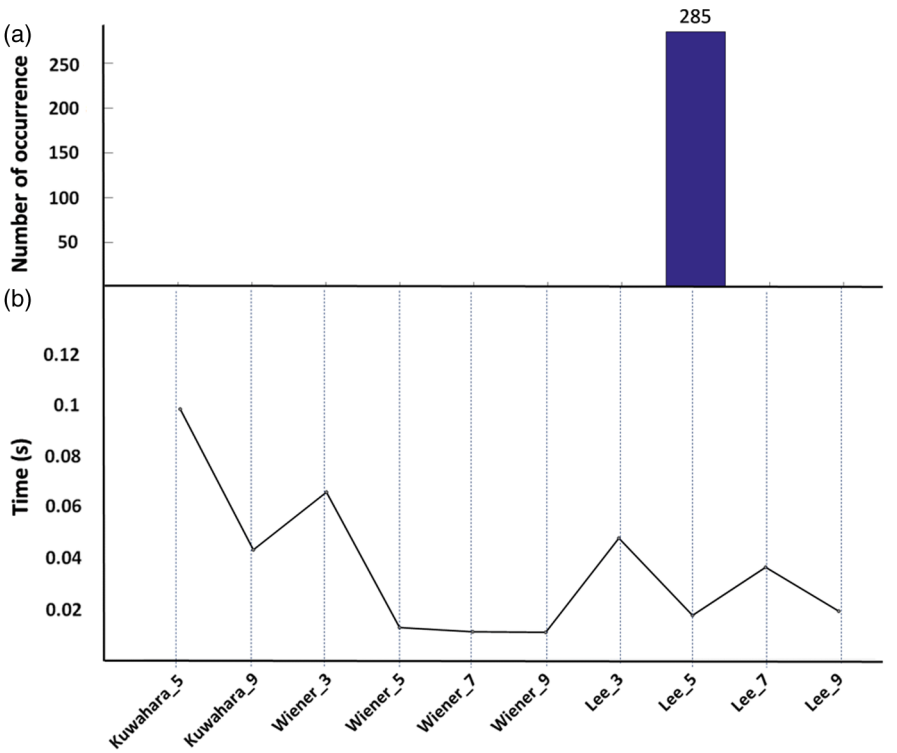

Fig. 7 Histogram of winner filter for 285 OCT images with their corresponding execution time. Adaptive statistical filters are used here. Histogram (a) illustrates the number of images that each filter is winning on and (b) shows the running time for each filtering method.

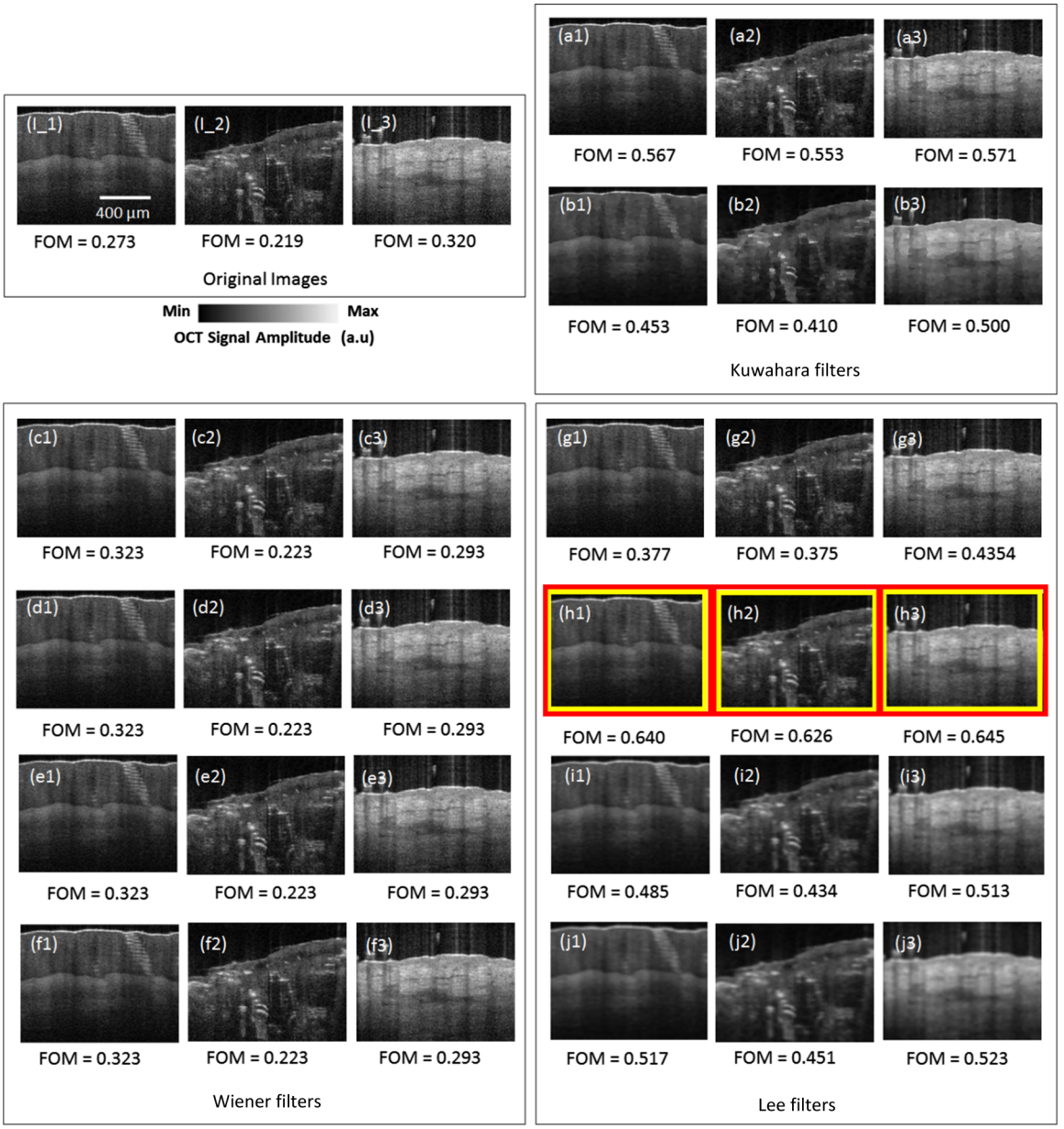

Fig. 8 Results of despeckling using adaptive statistical filters on OCT images used in Fig. 6 (I_1–I_3). (a–b) Despeckled images using Kuwahara filter with window sizes 5 and 9, respectively, (c–f) despeckled images using Wiener filter with window sizes 3, 5, 7, and 9, respectively, (g–j) despeckled images using Lee filters with window sizes 3, 5, 7, and 9, respectively. The yellow boxes indicate the winner filter based on FOM measure, the red boxes indicate the winner filter chosen by the classifier in adaptive statistical filters subgroup.

Figure 7 shows the histogram of winner filters in the adaptive statistical filter category for 285 OCT images as well as their execution time. According to this graph, for our image set, the Lee filter with window size 5 is chosen for all the images that were used in this study when FOM was the quality assessment criterion.

In Fig. 8, original OCT images and despeckled ones using adaptive statistical filters are shown. Here the classifier has only learned the adaptive statistical filtering subgroup. The yellow boxes in the figure indicate the winner filters chosen based on FOM criterion, i.e., here, Lee filter with window size 5. The red boxes indicate the winner filters chosen by the classifier in the adaptive statistical filtering subgroup.

Figure 9 shows the histogram of winner filters in the patch or pixel correlation filter category as well as their execution time. According to this graph, when FOM is the quality assessment criterion, in most cases BM3D filter is chosen as the winner filter. NLM and TV are the next winners, respectively.

In Fig. 10, original OCT images and despeckled ones using patch or pixel correlation filters are shown. In this case, the classifier has only learned the patch or pixel correlation filtering subgroup. The yellow boxes in the figure indicate the winner filters selected based on FOM criterion, i.e., BM3D and NLM. The red box indicates the winner filter chosen by classifier.

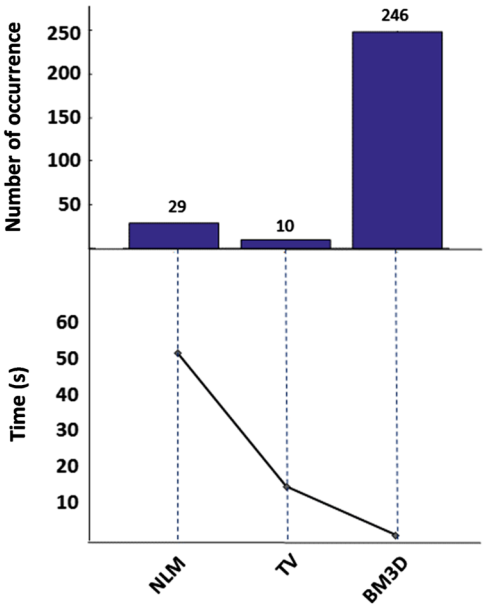

Fig. 9 Histogram of winner filter of 285 OCT images with their corresponding execution time. Patch or pixel correlation filters are used here. Histogram (a) illustrates the number of images that each filter is winning on and (b) shows the running time for each filtering method.

3.2 *Classifier Training and Performance Evaluation*

With the above experiments and the obtained results, we demonstrated that FOM is a reliable merit and could be used as the class label in the training of the classifier. This was shown by comparing the winner filter chosen based on FOM (calculated regardless of image textural features and relying only on quality assessment measures) and that chosen based on the classifier (trained based on image's features). We evaluated the selection rate of filters in all three categories together for the 285 images (see Fig. 11).

Figure 12 (I_1–I_6) shows the results of LDF on six OCT test images (a_1 to a_6) with their corresponding winner filter predicted by the classifier when all 25 filters are considered.

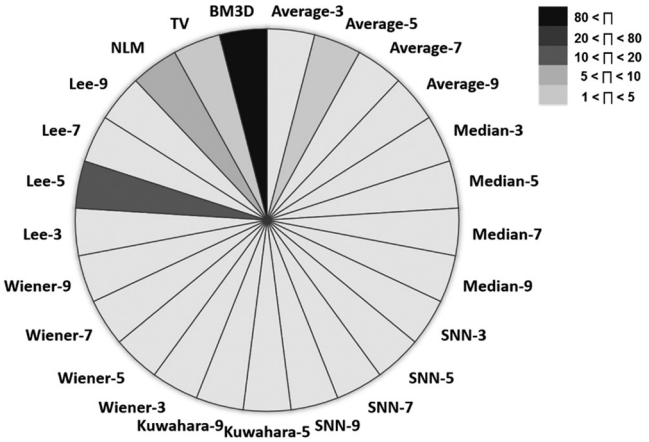

Fig. 11 The probability map (∏) of selecting each filter based on classifier's results for the 285 OCT images used in this study.

Figures 12(b)–12(d) show normalized FOM values for all 25 filters for the OCT images. Comparing the results of the winner filter chosen by classifier with that chosen by the normalized FOM, it shows that the classifier can predict the winner filter with high accuracy, without having to know the result of each individual filter and only based on the features extracted from the image. In this experiment, the classifier accuracy was measured as 97% based on a fivefold cross-validation method.[42]

Regarding the execution time, we observed that even though the patch or pixel correlation category filters filtered images most efficiently, their execution time is in the order of seconds, whereas the execution time of most of the sliding window filters are in millisecond range. The choice of having a better quality (based on FOM) or shorter execution time can be added to LDF criteria.

Speckle is an artifact that affects image quality and resolution of OCT images. Computer reduction of speckle has been a subject of interest since the early times of laser speckle and digital image processing,[2,3,44,45] and their application to OCT have led to a large number of publications over nearly 20 years. However, several groups including our group have studied different

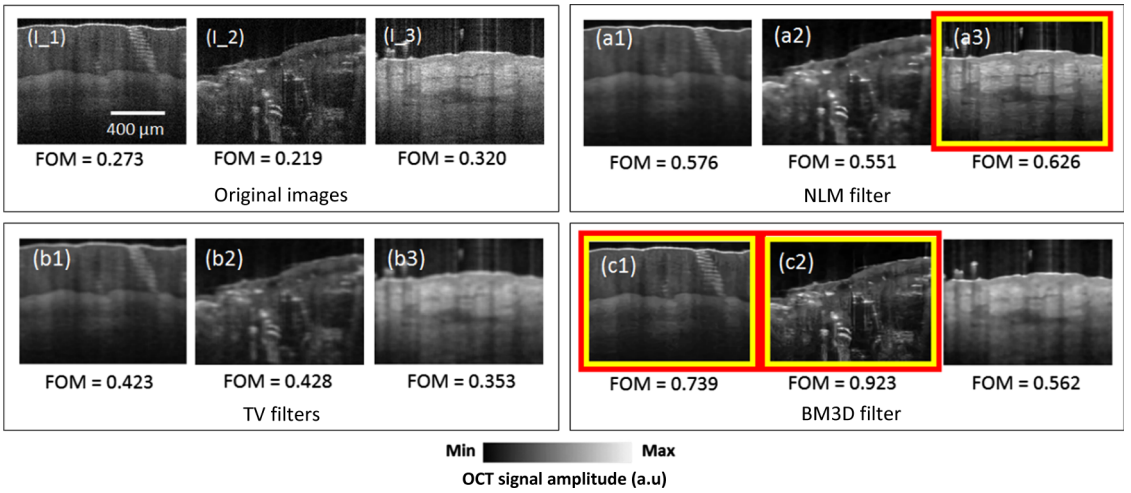

Fig. 10 Results of despeckling using adaptive statistical filters on OCT images used in Fig. 6 (I_1–I_3). (a) Despeckled images using NLM filter, (b) despeckled images using TV filter, (c) despeckled images using BM3D filter. The yellow boxes indicate the winner filters based on the FOM measure, the red boxes indicate the winner filter chosen by the classifier in patch or pixel correlation filters subgroup.

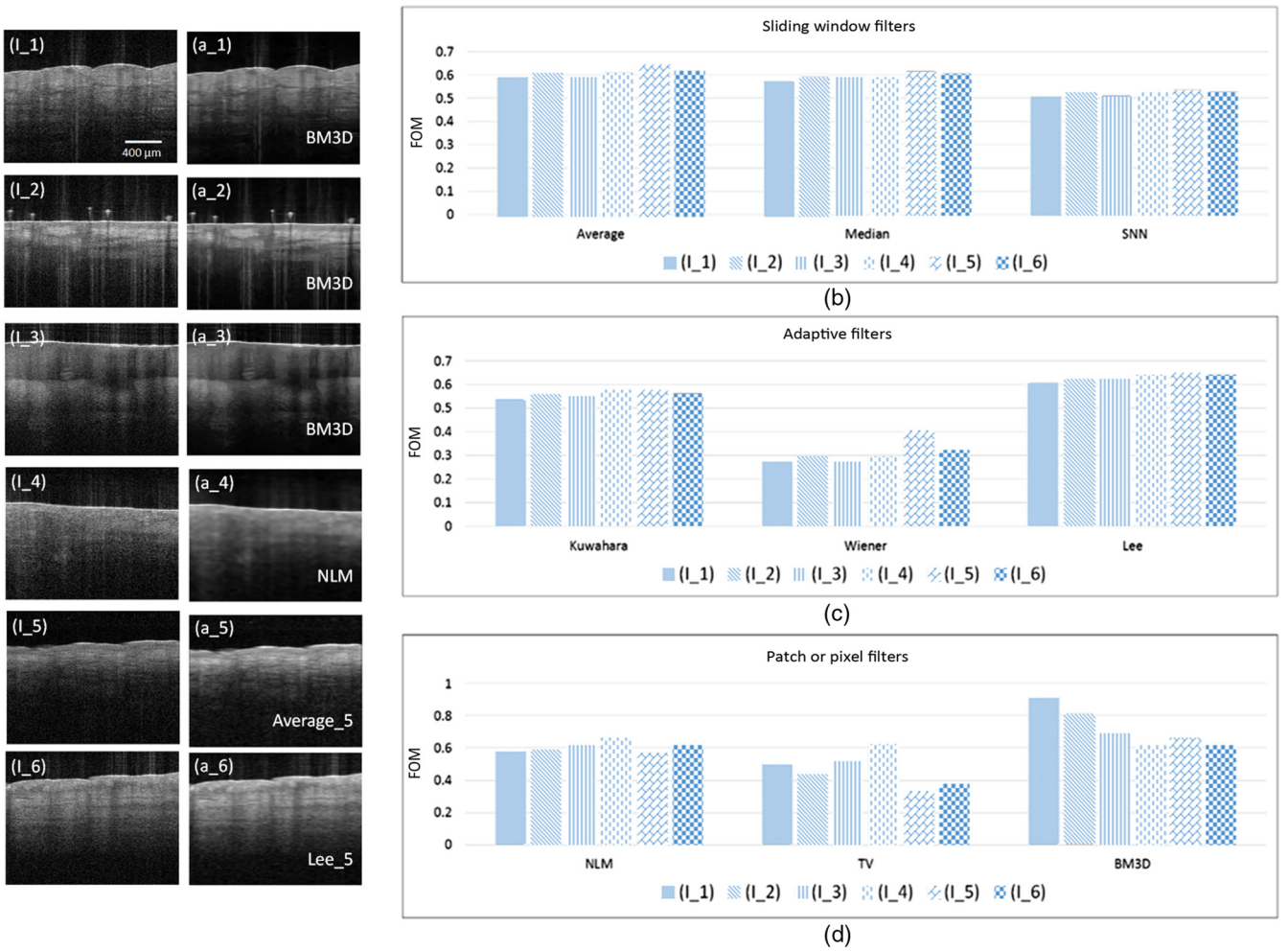

Fig. 12 (left) Results of LDF on six OCT test images with the corresponding winner filter predicted by the classifier. (I_1–I_6) Six original OCT images, and their winner filters, i.e., (a_1–a_3) despeckled images by BM3D filter as the selected optimum filter, (a_4) despeckled image by NLM filter as the selected optimum filter, (a_5) despeckled image using average filter with window sizes 5 as the selected optimum filter, (a_6) despeckled image using Lee filter with window sizes 5 as the selected optimum filter. (right) Normalized FOM for all the 25 filters for the same six OCT images for (b) sliding window filters, (c) adaptive filters, and (d) path- or pixel-based filters.

speckle reduction methods for OCT images, finding an optimum speckle reduction filter for an image or image set has not been comprehensively explored. We propose a procedure; we called it LDF, to learn the most effective speckle reduction method for one given set of images. We believe that LDF could help in choosing the most appropriate despeckling filter based on tissue morphological, textural, and optical features.

Using a more diverse and a larger number of OCT images (both healthy and diseased) for training the classifier, utilizing a more efficient training algorithm along with a larger number of features and quality assessment measures are planned as our future work.

4 Conclusion

Speckle is an artifact that affects image quality and resolution. Speckle "noise" reduction has been a subject of interest since spatially coherent lasers were invented. The necessity of such algorithms for OCT images has led to a large number of algorithms in the literature. We propose an expandable and learnable framework to organize these filters in an intelligent manner. The framework finds the most suitable speckle reduction algorithm for a given image. We called this framework LDF. LDF learns an FOM as a single quantitative image assessment measure constructed from SNR, CNR, SNL, EPI, and MSSIM using an autoencoder neural network. LDF was then trained to decide, which speckle reduction algorithm is most effective for a given image based on the image textural and optical features. The classification accuracy of LDF is determined as 97%. We tested LDF on 285 images. The quality of the despeckled images using LDF is not only improved in one quality metric but also in a combination of all quality metrics. LDF can also be customized based on a sole feature that the user requests. Utilizing a larger dataset for training the encoder and classifier as well as utilizing a more efficient training algorithm along with a larger number of uncorrelated features as well as quality assessment measures are planned as our future work.

Disclosures

The authors have no relevant financial interests in the article and no other potential conflicts of interest to disclose.

Acknowledgments

This project has been funded by Michelson Diagnostics and by Wayne State University Startup fund.

References

1. S. Adabi et al., "An overview of methods to mitigate artifacts in optical coherence tomography imaging of the skin," *Skin Res. Technol.* **1**, 1–9 (2017).
2. A. M. Zysk et al., "Optical coherence tomography: a review of clinical development from bench to bedside," *J. Biomed. Opt.* **12**(5), 051403 (2007).
3. W. Drexler et al., "Optical coherence tomography today: speed, contrast, and multimodality," *J. Biomed. Opt.* **19**(7), 071412 (2014).
4. J. M. Schmitt, "Optical coherence tomography (OCT): a review," *IEEE J. Sel. Top. Quantum Electron.* **5**(4), 1205–1215 (1999).
5. M. R. N. Avanaki et al., "Speckle reduction using an artificial neural network algorithm," *Appl. Opt.* **52**, 21, 5050–5057 (2013).
6. J. M. Schmitt, S. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.* **4**(1), 95–105 (1999).
7. T. M. Jørgensen et al., "Speckle reduction in optical coherence tomography images of human skin by a spatial diversity method," in *European Conf. on Biomedical Optics* (2007).
8. A. Ozcan et al., "Speckle reduction in optical coherence tomography images using digital filtering," *J. Opt. Soc. Am. A* **24**(7), 1901–1910 (2007).
9. D. C. Adler, T. H. Ko, and J. G. Fujimoto, "Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter," *Opt. Lett.* **29**(24), 2878–2880 (2004).
10. J.-S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Comput. Graphics Image Process.* **17**(1), 24–32 (1981).
11. J. E. Kyprianidis, H. Kang, and J. Döllner, "Image and video abstraction by anisotropic Kuwahara filtering," *Comput. Graphics Forum* **28**, 1955–1963 (2009).
12. J. S. Lim, *Two-Dimensional Signal and Image Processing*, p. 710, Prentice Hall, Englewood Cliffs, New Jersey (1990).
13. P. Puvanathasan and K. Bizheva, "Speckle noise reduction algorithm for optical coherence tomography based on interval type II fuzzy set," *Opt. Express* **15**(24), 15747–15758 (2007).
14. J. Cheng et al., "Speckle reduction in 3D optical coherence tomography of retina by A-scan reconstruction," *IEEE Trans. Med. Imaging* **35**(10), 2270–2279 (2016).
15. A. Wong et al., "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery," *Opt. Express* **18**(8), 8338–8352 (2010).
16. D. Harwood et al., "A new class of edge-preserving smoothing filters," *Pattern Recognit. Lett.* **6**(3), 155–162 (1987).
17. A. Uzan, Y. Rivenson, and A. Stern, "Speckle denoising in digital holography by nonlocal means filtering," *Appl. Opt.* **52**(1), A195–A200 (2013).
18. D. P. Popescu et al., "Signal attenuation and box-counting fractal analysis of optical coherence tomography images of arterial tissue," *Biomed. Opt. Express* **1**(1), 268–277 (2010).
19. Y. Chen et al., "Effects of axial resolution improvement on optical coherence tomography (OCT) imaging of gastrointestinal tissues," *Opt. Express* **16**(4), 2469–2485 (2008).
20. P. B. Garcia-Allende et al., "Morphological analysis of optical coherence tomography images for automated classification of gastrointestinal tissues," *Biomed. Opt. Express* **2**(10), 2821–2836 (2011).
21. P. M. Shankar, "Speckle reduction in ultrasound B-scans using weighted averaging in spatial compounding," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **33**(6), 754–758 (1986).
22. J. Rogowska and M. E. Brezinski, "Evaluation of the adaptive speckle suppression filter for coronary optical coherence tomography imaging," *IEEE Trans. Med. Imaging* **19**(12), 1261–1266 (2000).
23. D. L. Marks, T. S. Ralston, and S. A. Boppart, "Speckle reduction by I-divergence regularization in optical coherence tomography," *J. Opt. Soc. Am. A* **22**(11), 2366–2371 (2005).
24. P. Bakker, L. J. van Vliet, and P. W. Verbeek, "Edge preserving orientation adaptive filtering," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition* (1999).
25. A. H. Samak, "A new nonlinear anisotropic-wiener method for speckle noise reduction in optical coherence tomography," *Int. J. Comput. Appl.* **65**(12) (2013).
26. A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2005)*, pp. 60–65 (2005).
27. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
28. A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail-preserving filters for image processing," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**(1), 74–90 (1987).
29. A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE Trans. Geosci. Remote Sens.* **28**(6), 992–1000 (1990).
30. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D* **60**(1), 259–268 (1992).
31. K. Dabov et al., "Image restoration by sparse 3D transform-domain collaborative filtering," *Proc. SPIE* **6812**, 681207 (2008).
32. K. W. Gossage et al., "Texture analysis of optical coherence tomography images: feasibility for tissue classification," *J. Biomed. Opt.* **8**(3), 570–575 (2003).
33. K. W. Gossage et al., "Texture analysis of speckle in optical coherence tomography images of tissue phantoms," *Phys. Med. Biol.* **51**(6), 1563–1575 (2006).
34. R. M. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.* **SMC-3**(6), 610–621 (1973).
35. M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graphics Image Process.* **4**(2), 172–179 (1975).
36. K. Vermeer et al., "Depth-resolved model-based reconstruction of attenuation coefficients in optical coherence tomography," *Biomed. Opt. Express* **5**(1), 322–337 (2014).
37. J. O. Ramsay, *Functional Data Analysis*, Wiley Online Library, Indianapolis (2006).
38. H. M. Salinas and D. C. Fernandez, "Comparison of PDE-based nonlinear diffusion approaches for image enhancement and denoising in optical coherence tomography," *IEEE Trans. Med. Imaging* **26**(6), 761–771 (2007).
39. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
40. M. R. Avanaki et al., "Spatial compounding algorithm for speckle reduction of dynamic focus OCT images," *IEEE Photonics Technol. Lett.* **25**(15), 1439–1442 (2013).
41. G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science* **313**(5786), 504–507 (2006).
42. M. R. N. Avanaki and A. Podoleanu, "En-face time-domain optical coherence tomography with dynamic focus for high-resolution imaging," *J. Biomed. Opt.* **22**, 5, 056009 (2017).
43. P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of Database Systems*, pp. 532–538, Springer (2009).
44. A. Hojjatoleslami and M. R. N. Avanaki, "OCT skin image enhancement through attenuation compensation," *Appl. Opt.* **51**(21), 4927–4935 (2012).
45. S. A. Hojjatoleslami, M. R. N. Avanaki, and A. G. Podoleanu, "Image quality improvement in optical coherence tomography using Lucy-Richardson deconvolution algorithm," *Appl. Opt.* **52**(23), 5663–5670 (2013).

Biographies for the authors are not available.