

# Adaptive compounding speckle-noise-reduction filter for optical coherence tomography images

Juan J. Gómez-Valverde<sup>a,b,\*</sup>, Christoph Sinz,<sup>c</sup> Elisabet A. Rank<sup>d</sup>,  
Zhe Chen,<sup>d</sup> Andrés Santos<sup>a,b</sup>, Wolfgang Drexler,<sup>d</sup> and  
María J. Ledesma-Carbayo<sup>a,b</sup>

<sup>a</sup>Universidad Politécnica de Madrid, ETSI Telecomunicación,  
Biomedical Image Technologies Laboratory, Madrid, Spain

<sup>b</sup>Biomedical Research Center in Bioengineering, Biomaterials and Nanomedicine,  
Madrid, Spain

<sup>c</sup>Medical University of Vienna, Department of Dermatology, Vienna, Austria

<sup>d</sup>Medical University of Vienna, Center for Medical Physics and Biomedical Engineering,  
Vienna, Austria

## Abstract

**Significance:** Speckle noise limits the diagnostic capabilities of optical coherence tomography (OCT) images, causing both a reduction in contrast and a less accurate assessment of the micro-structural morphology of the tissue.

**Aim:** We present a speckle-noise reduction method for OCT volumes that exploits the advantages of adaptive-noise wavelet thresholding with a wavelet compounding method applied to several frames acquired from consecutive positions. The method takes advantage of the wavelet representation of the speckle statistics, calculated properly from a homogeneous sample or a region of the noisy volume.

**Approach:** The proposed method was first compared quantitatively with different state-of-the-art approaches by being applied to three different clinical dermatological OCT volumes with three different OCT settings. The method was also applied to a public retinal spectral-domain OCT dataset to demonstrate its applicability to different imaging modalities.

**Results:** The results based on four different metrics demonstrate that the proposed method achieved the best performance among the tested techniques in suppressing noise and preserving structural information.

**Conclusions:** The proposed OCT denoising technique has the potential to adapt to different image OCT settings and noise environments and to improve image quality prior to clinical diagnosis based on visual assessment.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.26.6.065001](https://doi.org/10.1117/1.JBO.26.6.065001)]

**Keywords:** optical coherence tomography; denoising; image processing; speckle; wavelets.

Paper 210051R received Feb. 17, 2021; accepted for publication May 24, 2021; published online Jun. 17, 2021.

## 1 Introduction

Optical coherence tomography (OCT)<sup>1</sup> is an optical imaging technique that allows cross-sectional views of *in vivo* tissue in real time with micrometer resolution and at depths of up to two millimeters. OCT is widely used in a variety of biomedical and clinical fields, such as ophthalmology (as a routine noninvasive tool for the diagnosis and monitoring of disease progression)<sup>2</sup> and cardiology (as a catheter-based imaging system during coronary intervention).<sup>3</sup> In dermatology,

\*Address all correspondence to Juan J. Gómez-Valverde, [juanjo.gomez@upm.es](mailto:juanjo.gomez@upm.es)

OCT has been used to study a variety of dermatological disorders.<sup>4,5</sup> The most significant dermatological applications of OCT are in assessing non-melanoma skin cancers such as basal cell carcinoma (BCC) and in reducing the need for diagnostic biopsies.<sup>6-8</sup>

Because it uses spatially coherent illumination, OCT images are affected by the speckle phenomenon, which has a dual role as both a source of noise and a carrier of information.<sup>9,10</sup> Speckle carrying signal information is the result of the back scattering of the incident photons, whereas speckle noise is caused by the random interference between multiple reflected photons coming from multiple directions. Speckle noise gives a grainy appearance to the OCT images, which degrades the signal-to-noise ratio (SNR) and limits the accuracy of its interpretation. The speckle properties are affected by the scale representation, the optical settings, and the scattering properties of the biological tissue.<sup>11</sup>

Speckle denoising is an active and widespread research field developed during recent years. There are two main approaches: those that modify the image configuration (where the optical setting or the scanning protocol can be adjusted) and those that are based on postprocessing the images via digital algorithms. In the first group, there are three main approaches to the compounding techniques: using the frequency,<sup>12</sup> using the angle of the incident light source,<sup>13</sup> and using multiple A-lines collected in a controlled way.<sup>14</sup> Other studies have proposed methods for combining image processing and modifications to the optical configuration, with the aim of increasing the SNR even further.<sup>15</sup> To compensate for the effect of possible movement in the case of *in vivo* imaging, two strategies for reducing motion artifacts have been proposed: using a weighted average of individual A-scans<sup>16</sup> and using a combination of the average of B-scans (from a high-speed CMOS line-scan camera) and a cross correlation of different frames with respect to a fixed reference. One recent report suggests that SNR improvements via pure angular compounding techniques will be limited by optical aberrations.<sup>17</sup> To overcome these limitations, a recent study proposed the combination of angular compounding with geometric image registration and digital focusing.<sup>18</sup>

The main advantage of digital speckle-reduction techniques is that these can be applied to almost all two-dimensional (2D) and three-dimensional (3D) images acquired by an OCT device without changing the acquisition setup. However, they usually add to the computation requirements and can affect the resolution of the image. These techniques can involve the combined use of several methods such as local averaging over neighboring A-scans of each tomogram,<sup>19</sup> averaging multiple B-scans,<sup>16</sup> applying rotation kernel transformations to each tomogram,<sup>20</sup> image regularization,<sup>21</sup> complex diffusion filtering,<sup>22</sup> curvelet transforms,<sup>23</sup> applying sparse representations or low rank models using patches in the images,<sup>24-27</sup> autoencoding based on a neural network that learns from a figure of merit,<sup>28</sup> digital filtering clusters of pixels with similar optical properties,<sup>29</sup> and adaptive nonlinear diffusion filtering.<sup>30</sup> State-of-the-art general-purpose denoising filters such as probability-based non-local-means (PNLM) and block-matching 3D filtering have also been adapted successfully for the removal of noise from OCT images.<sup>31-33</sup> Recently, approaches based on deep learning, such as denoising convolutional neural network (DNCNN),<sup>34</sup> have been proposed for natural image denoising and for speckle noise in OCT,<sup>35-39</sup> showing the potential of these techniques in a variety of OCT image types.

Algorithms based on filtering in the wavelet domain<sup>40</sup> have shown excellent performance in speckle noise removal. One approach used in these methods is to filter the detail wavelet coefficients in multiple subbands to minimize the noise. The calculation of an appropriate threshold and its application can be performed using spatially adaptive soft-thresholding with estimation of the noise in one subband.<sup>41,42</sup> However, a recent study shows that speckle noise can have different magnitudes for different wavelet subbands and that estimation of the noise variance at individual scales can improve the characterization of the threshold and can improve the speckle noise removal.<sup>43</sup> Another successful strategy is the compounding of several frames in combination with digital filtering,<sup>44</sup> more specifically with frames previously filtered using wavelet denoising.<sup>45</sup> The use of volumetric data collected from consecutive B-scans has also been exploited by other methods<sup>23,46,47</sup> as an additional source of information for use in speckle reduction.

Here, we present a method for speckle reduction of OCT images that effectively combines an adaptive noise strategy with volumetric wavelet compounding. The method first processes several frames acquired at consecutive locations, using a multiscale noise adaptive wavelet filter,

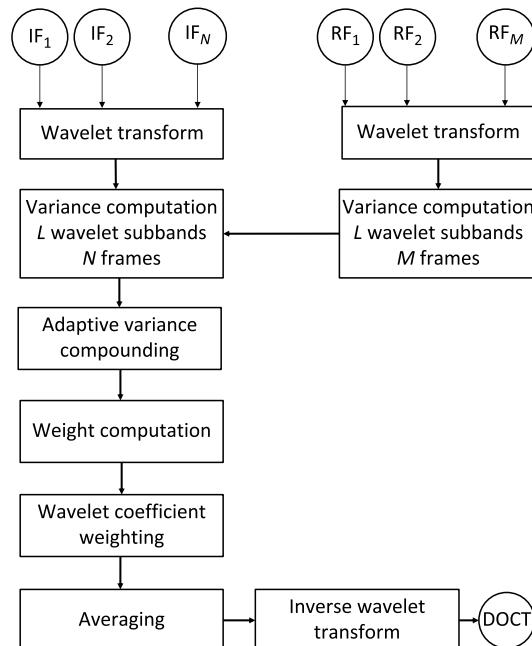
followed by an appropriately weighted computation for the compounding. Following Zaki et al.,<sup>43</sup> we estimate the noise-variance wavelet representation in a homogeneous scattering sample and use it to filter the noise in all frames. To take advantage of complementary information brought from several acquisitions, we used Mayer et al.'s<sup>45</sup> compounding approach as a second step, applying a weighting method to select the detail coefficients of each subband before compounding all of the filtered frames. Our method also benefits from the fact that the wavelet representation of the speckle statistics, calculated properly from a homogeneous sample or a region of the noisy volume previously recorded, can be used to characterize the noise pattern of the system. These reference statistics are a valuable asset when differentiating information from speckle noise. In addition, the use of several frames and the wavelet weighted compounding improves speckle removal and enhances the structural details, resulting in a superior performance to that of other state-of-the-art methods.

The rest of the paper is structured in several sections. In Sec. 2, we describe in detail the proposed method, the OCT imaging systems used, the datasets, the performance metrics, and the quantitative evaluation process used in the assessment. In Sec. 3, we report the results and discussion of the parameter assessment and the quantitative evaluation of the algorithm. Finally, Sec. 4 contains the conclusions of the study.

## 2 Materials and Methods

### 2.1 Wavelet-Compounding Adaptive Noise Filter

Figure 1 shows a complete overview of the proposed wavelet-compounding adaptive noise filter (WCAN) method. The processing steps are wavelet transform, variance computation of the input and reference frames, adaptive variance compounding, weight computation, wavelet coefficient



**Fig. 1** Flow diagram of the wavelet-compounding adaptive noise algorithm. The acquired OCT input frames  $IF_1, IF_2, \dots, IF_N$  represent the  $N$  input noisy images for the algorithm. All images and the  $M$  reference OCT frames  $RF_1, RF_2, \dots, RF_M$  are wavelet decomposed in  $L$  subbands. The variance of all subbands in the wavelet decomposition of the reference stack is computed. Using the subband variances, the weights of the detail coefficients for each subband are calculated based on the compounding of the previous variance computation and applied to each original detail coefficient. All denoised coefficients are then averaged in the wavelet domain before the inverse wavelet transform is calculated to obtain the final denoised OCT image, DOCT.

weighting, averaging, and inverse wavelet transform. The input to the algorithm is a set of  $N$  OCT frames ( $\text{IF}_1, \text{IF}_2, \dots, \text{IF}_N$ ) and a set of  $M$  reference OCT frames ( $\text{RF}_1, \text{RF}_2, \dots, \text{RF}_M$ ), with both sets acquired at consecutive positions. The reference frames could come from a homogeneous scattering sample independent of the input frames or from a homogeneous region of the set of input frames. This reference set should be collected as a prior step to the application of the method. The output of the algorithm is a single OCT denoised image. The terms “frame” and “image” from one side and the terms “method,” “filter,” and “algorithm” are used interchangeably in this paper. In the following paragraphs, we explain in detail each step of the method.

### 2.1.1 OCT image noise

The speckle noise is usually modeled as multiplicative due to the multiple backscattering effects.<sup>48</sup> In addition to the speckle noise, OCT systems are also affected by noise coming from different sources such as inherent laser intensity noise, photonics shot noise, or thermal noise from electronics. These other types can be assumed to be white Gaussian additive noise.<sup>49</sup> We define an OCT image  $I^*$  as

$$I^* = N_s \cdot I + N_w, \quad (1)$$

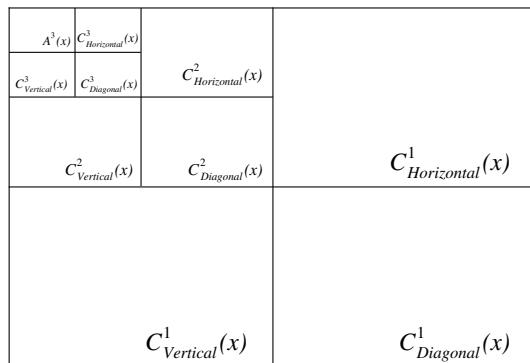
where  $I$  is the noise-free image and  $N_s$  and  $N_w$  are the speckle noise and the Gaussian white noise, respectively. We can neglect the additive white noise because it is significantly small compared with the multiplicative noise.<sup>50</sup> Thus, we model the OCT image as

$$I^* \approx N_s \cdot I. \quad (2)$$

To convert the multiplicative noise into additive noise, we consider that all of the OCT images are logarithmic transformed following the approach in similar studies.<sup>50-54</sup> This operation allows us to assume a near additive noise model after the logarithmic transformation. We assume also that speckle noise present in frames recorded at different consecutive positions is uncorrelated.<sup>45</sup>

### 2.1.2 Wavelet transform

All of the input images are decomposed by a wavelet transformation with a maximum decomposition level of  $L$ . The result of the transformation is a set of approximation coefficients  $A_i^l$  and detail coefficients  $C_{i,O}^l$ , where  $i$  is the frame number,  $l$  is the decomposition level, and  $O$  is the orientation or direction of the detail coefficient (horizontal, vertical, or diagonal). The wavelet transformation is computed using the 2D discrete stationary wavelet transform.<sup>40</sup> Figure 2 shows



**Fig. 2** Representation of a 2D wavelet decomposition of an image using three levels.

a representation of a 2D wavelet decomposition of an image with three levels, where  $A$  contains the approximation coefficients and  $C_{\text{Horizontal}}$ ,  $C_{\text{Vertical}}$ , and  $C_{\text{Diagonal}}$  are the detail coefficients of each subband for the horizontal, vertical, and diagonal orientation, respectively.

### 2.1.3 Variance computation

The next step is the calculation of the variance of the detail coefficients of all of the subbands of the input and reference frames. For the computation of the variance of the detail coefficients of each subband  $\sigma_O^{l2}$ , we use the following expressions:

$$\sigma_O^{l2} = \frac{1}{X-1} \sum_{x=1}^X |C_O^l(x) - \mu_O^l|^2, \quad (3)$$

$$\mu_O^l = \frac{1}{X} \sum_{k=1}^X C_O^l(k), \quad (4)$$

where  $l$  is the subband level (between 1 and  $L$ ),  $O$  is the orientation (horizontal, vertical, or diagonal) of the detail coefficient  $C$ , and  $x$  is an index ranging from 1 to  $X$  (total number of coefficients with orientation  $O$  for the subband  $l$ ). Equation (3) uses  $\mu_O^l$ , the mean of all detail coefficients for the orientation  $O$  and subband  $l$ , calculated using Eq. (4). In the case of the reference frames, we calculate the variance,  $\sigma_{O,R}^{l2}$ , considering all  $M$  frames, and therefore,  $X$  ranging in this case, from 1 to the number of coefficients by subband and by orientation plus the number of frames  $M$ .

### 2.1.4 Adaptive variance compounding

Then, we calculate the adaptive compounding terms  $G_{O,i}^l$ , which consider an average of the ratio of the variance terms of the initial frames and the reference frames, including also the detail coefficients for the variance of the frame  $i$ , using the following expressions:

$$G_{O,i}^l(x) = \frac{1}{N} \left( |C_{O,i}^l(x)| \frac{\sigma_{O,i}^{l2}}{\sigma_{O,R}^{l2}} + \sum_{j=1 \wedge j \neq i}^{N-1} \frac{\sigma_{O,j}^{l2}}{\sigma_{O,R}^{l2}} \right), \quad (5)$$

where  $l$  is the subband,  $i$  is the frame number ranging from 1 to  $N$  (total number of input frames, see Fig. 1) and  $O$  is the orientation of the detail coefficient (horizontal, vertical, or diagonal). The term  $C_{O,i}^l$  is the detail coefficient of the subband  $l$ , orientation  $O$ , and frame  $i$ . The terms  $\sigma_{O,i}^{l2}$  and  $\sigma_{O,R}^{l2}$  are the variance of the detail coefficients from the frame  $i$  and from the reference frames respectively, calculated using Eq. (3).

### 2.1.5 Weight computation

The weights provide estimates of the noise contribution of each subband to every frame in relation to the reference noise and to the remaining frames. The goal is to reduce the detail coefficients in each subband with higher variance compared with the reference values and the remaining frames, assuming that noise will be the main cause of this higher contribution. The term in Eq. 6 is an adaptive variance threshold defined in a similar way to standard wavelet thresholding, but considering the variance of the reference frames in each subband.

$$T_{O,i}^l = \frac{\sigma_{O,R}^{l2}}{\sqrt{|\sigma_{O,i}^{l2} - \sigma_{O,R}^{l2}|}}, \quad (6)$$

$$W_{O,i}^l(x) = \begin{cases} 1, & |C_{O,i}^l(x)| > k \cdot T_{O,i}^l \\ 1 - G_{O,i}^l(x), & \text{otherwise} \end{cases} \quad (7)$$

Equation (7) produces the final weights  $W_{O,i}^l$  for each subband  $l$ , frame  $i$ , and orientation  $O$ , comparing the value of the detail coefficients  $C_{O,i}^l$  with the previous threshold value  $T_{O,i}^l$ . The parameter  $k$  will then be used to balance the final amount of noise reduction (in Sec. 3.1, we present more detail of the influence and utility of this parameter). The ratio  $G_{O,i}^l$  from Eq. (5) is used to reduce the weight of the detail coefficients lower than the threshold  $T_{O,i}^l$ .

### 2.1.6 Wavelet coefficient weighting, averaging, and inverse wavelet transform

The new detail coefficients  $\tilde{C}_{O,i}^l$  of each subband  $l$ , orientation  $O$ , and initial frame  $i$  are computed using the previous weights for all positions  $x$ , via the following expression:

$$\tilde{C}_{O,i}^l(x) = C_{O,i}^l(x) \cdot W_{O,i}^l(x). \quad (8)$$

The detail  $C_O^l$  and the approximation  $A^l$  coefficients of the denoised image (see Fig. 2) for each subband  $l$  are calculated by averaging the coefficients of the  $N$  initial frames considered during the process:

$$C_O^l(x) = \frac{1}{N} \sum_{i=1}^N \tilde{C}_{O,i}^l(x), \quad (9)$$

$$A^l(x) = \frac{1}{N} \sum_{i=1}^N A_i^l(x). \quad (10)$$

Finally, the denoised image is computed by applying the inverse wavelet transform over the averaged coefficients. The algorithm was implemented in Matlab 2018b (MathWorks, Inc., Natick, Massachusetts) on a personal computer (Intel 3.3 GHz CPU, 32 GB memory).

### 2.1.7 Data and OCT imaging systems

For the quantitative evaluation of the method, we used datasets representing four different environments, as specified in Table 1. Datasets Medical University of Vienna (MUW)-1, MUW-2, and MUW-3 each comprised 18 skin OCT volumes and were acquired via three different OCT

**Table 1** Technical specifications of the OCT imaging systems used in the acquisition of the volumes in this study.

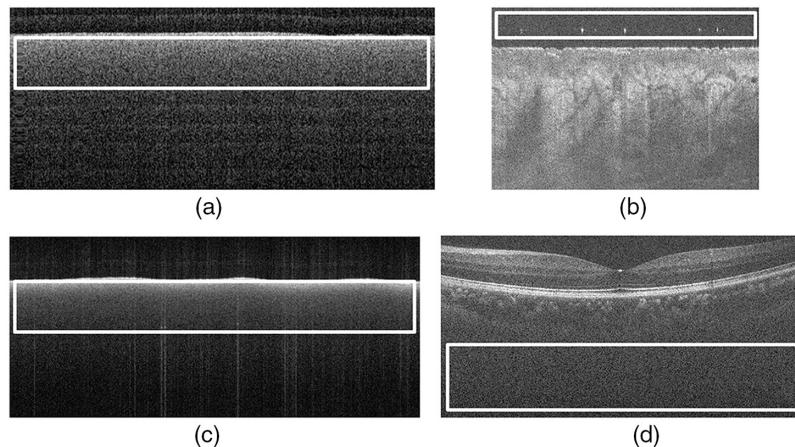
Parameter	MUW-1 <sup>55,56</sup>	MUW-2 <sup>57</sup>	MUW-3 (VivoSight)	A2ASDOCT (Bioptigen) <sup>58</sup>
System type	Swept-source frequency domain OCT	Spectral domain OCT	Multi-beam swept-source frequency domain OCT	Spectral domain OCT
Laser system	1340 nm	1320 nm	1305 nm Class 1	840 nm
A-line rate	200 kHz	47 kHz	10 kHz	17 kHz
Optical resolution	19.5 $\mu\text{m}$ (air) lateral 4.6 $\mu\text{m}$ (air) axial	15 $\mu\text{m}$ (lateral) 7 $\mu\text{m}$ (axial)	<7.5 $\mu\text{m}$ (tissue) lateral <5 $\mu\text{m}$ (tissue) axial	10 $\mu\text{m}$ (tissue) lateral 4.5 $\mu\text{m}$ (tissue) axial
Field of view	10 $\times$ 10 mm	7 $\times$ 3.5 mm	6 $\times$ 6 mm	6.7 $\times$ 6.7 mm
Imaging depth	1.2 mm	2 mm	Between 1.2 and 2 mm, tissue dependent	2 mm

**Table 2** Description of the datasets. The volumes of the three dermatological datasets were acquired and evaluated by experts from the Department of Dermatology at the MUW. The identified sample area (if available) is given in the fourth column. The results of the clinical assessments are given in the fifth column.

Dataset	Size (pixels)/number of B-scans	Volume ID	Sample area	Clinical evaluation
MUW-1	1100 × 512/256	1.1	Hand	Nevus araneus
		1.2	Back	Postsurgical scar
		1.3	Thigh	Normal skin
		1.4	Head	BCC
		1.5	Arm	BCC
MUW-2	1000 × 680/256	2.1	Palm	Normal skin
		2.2	—	Pityriasis
		2.3	Forearm	Normal skin
		2.4	Head	BCC
		2.5	Arm	Nevus
MUW-3	1558 × 460/50	3.1	Nose	BCC
		3.2	Chest	BCC
		3.3	Chest	BCC
		3.4	Palm	Normal skin
	1298 × 460/50	3.5	Fingertip	Normal skin
		3.6	Back	Nevus
		3.7	Abdomen	Angioma
		3.8	Cheek	Folliculitis
A2ASDOCT	900 × 450/5	4.1–4.17	Fovea	10 from normal subjects and 7 from AMD subjects

image systems. MUW-1 and MUW-2 were custom designs from the Center for Medical Physics and Biomedical Engineering at the MUW. To assess the applicability of the method to a commercial device, the MUW-3 image dataset was acquired via a VivoSight OCT scanner (Michelson Diagnostics, Kent). The A2ASDOCT dataset is a public dataset comprising 17 retinal volumes from the A2ASDOCT study.<sup>59</sup> It was produced using a Bioptigen Inc (Research Triangle Park, North Carolina) spectral domain OCT imaging system<sup>24</sup> and involved 17 eyes from 17 subjects with and without nonneovascular age-related macular degeneration (AMD). Each set of images in this dataset comprises five frames of 900 × 450 pixels from consecutive positions, each of which includes the fovea region. The main technical specifications of the imaging systems are summarized in Table 1. Table 2 gives additional details about the dermatological datasets (MUW-1, MUW-2, and MUW-3) and the retinal dataset (A2ASDOCT).

The method requires the use of a reference set ( $RF_1, RF_2, \dots, RF_M$  in Fig. 1). The role of this reference is to characterize the speckle noise assuming, as stated by Zaki et al.,<sup>43</sup> that the variation in these uniform areas is determined by random noise. The calculation of the noise variance is based on Eq. 3 over a region of interest (ROI) selected from the reference frames. We used two strategies to calculate this variance. First, we used a homogeneous scattering phantom made of synthetic clay (Blu-Tak®; Bostik, Wauwatosa, Wisconsin). This approach was tested with the MUW-1 OCT image system and the MUW-3 VivoSight device. Using the MUW-1 system,



**Fig. 3** Examples of frames and ROIs (white rectangles) used to compute the noise variance required by the WCAN method. (a) First frame with size  $512 \times 219$  pixels for the phantom acquired via the MUW-1 image system. (b) First frame with size  $1000 \times 680$  pixels for Volume 2.1 (palm normal skin, see Table 2) acquired via the MUW-2 image system. (c) First frame with size  $1038 \times 460$  pixels for the phantom acquired via the VivoSight device. (d) First frame with size  $900 \times 450$  pixels for the first set in the A2ASDOCT dataset.

we acquired 512 -scans with dimensions of  $512 \times 219$  pixels from consecutive positions. Using the VivoSight device, we acquired 120 B-scans with dimensions of  $1038 \times 460$  pixels from consecutive positions.

In our second strategy, we selected a ROI in the same location (top or bottom) of the frames from both systems and computed the variances of the coefficients of four wavelet detail subbands. This involved selecting a homogeneous noisy ROI and computing the corresponding variance of the detail coefficients, considering all frames in the dataset. We selected the location of the ROI (top or bottom) to have enough data in the all frames of the volume. We used this strategy as an alternative approach for the MUW-2 and A2ASDOCT datasets, partly because a phantom was unavailable and partly to test an alternative approach to noise estimation. Figure 3 shows examples of the acquisition of the phantom and the selection of the ROIs for both strategies, and Table 3 summarizes the reference sets used for each OCT imaging system. We applied the same reference set to all volumes of each dataset (MUW-1, MUW-2, MUW-3, and A2ASDOCT).

## 2.2 Quantitative evaluation

We evaluated the efficacy of the algorithm, using common speckle-reduction performance metrics<sup>41,43,44</sup> via a comparison with different state-of-the-art methods. These quantitative metrics were the *SNR*, the contrast-to-noise ratio (*CNR*), and the equivalent number of looks (*ENL*), as expressed below.

**Table 3** Summary of the reference sets used in the application of the method WCAN for all of the datasets.  $M$  indicates the maximum number of frames in the reference set ( $RF_1, RF_2 \dots RF_M$ ). The ROI Dimensions column refers to the size of the white rectangles shown in Fig. 3.

Dataset	Volume ID	Strategy	M	ROI dimensions	Figure
MUW-1	R.1	ROI phantom	512	$63 \times 478$	Fig. 3a
MUW-2	R.2	ROI Vol 2.1	256	$112 \times 993$	Fig. 3b
MUW-3	R.3	ROI phantom	120	$120 \times 1009$	Fig. 3c
A2ASDOCT	R.4	ROI all frames	85	$153 \times 885$	Fig. 3d

$$SNR = 10 \log_{10} \left( \frac{\mu_I^2}{\sigma_I^2} \right), \quad (11)$$

$$CNR = \frac{1}{R} \sum_{r=1}^R \frac{\mu_r - \mu_b}{\sqrt{\sigma_r^2 + \sigma_b^2}}, \quad (12)$$

$$ENL = \frac{1}{H} \sum_{h=1}^H \frac{\mu_h^2}{\sigma_h^2}. \quad (13)$$

*SNR* is defined in Eq. (11), where  $\mu_I$  indicates the mean of the OCT image and  $\sigma_I^2$  refers to the noise variance. In the definition of *CNR* [Eq. (12)],  $\mu_b$  and  $\sigma_b^2$  are the mean and variance for a background noise region, respectively, with  $\mu_r$  and  $\sigma_r^2$  being the mean and variance for all ROIs, respectively, including homogeneous and heterogeneous regions. *ENL* is a measure of the smoothness of a homogeneous ROI. In Eq (13),  $\mu_h$  and  $\sigma_h^2$  are the mean and variance of all  $H$  homogeneous ROIs, respectively. Except for the *SNR* calculations, the remaining parameters were computed from the logarithmic OCT images.

To assess the overall influence of the quantitative metrics, we include two additional figure-of-merits (*FOMs*):<sup>44</sup>

$$FOM_{SUM} = SNR_{Norm} + CNR_{Norm} + ENL_{Norm}, \quad (14)$$

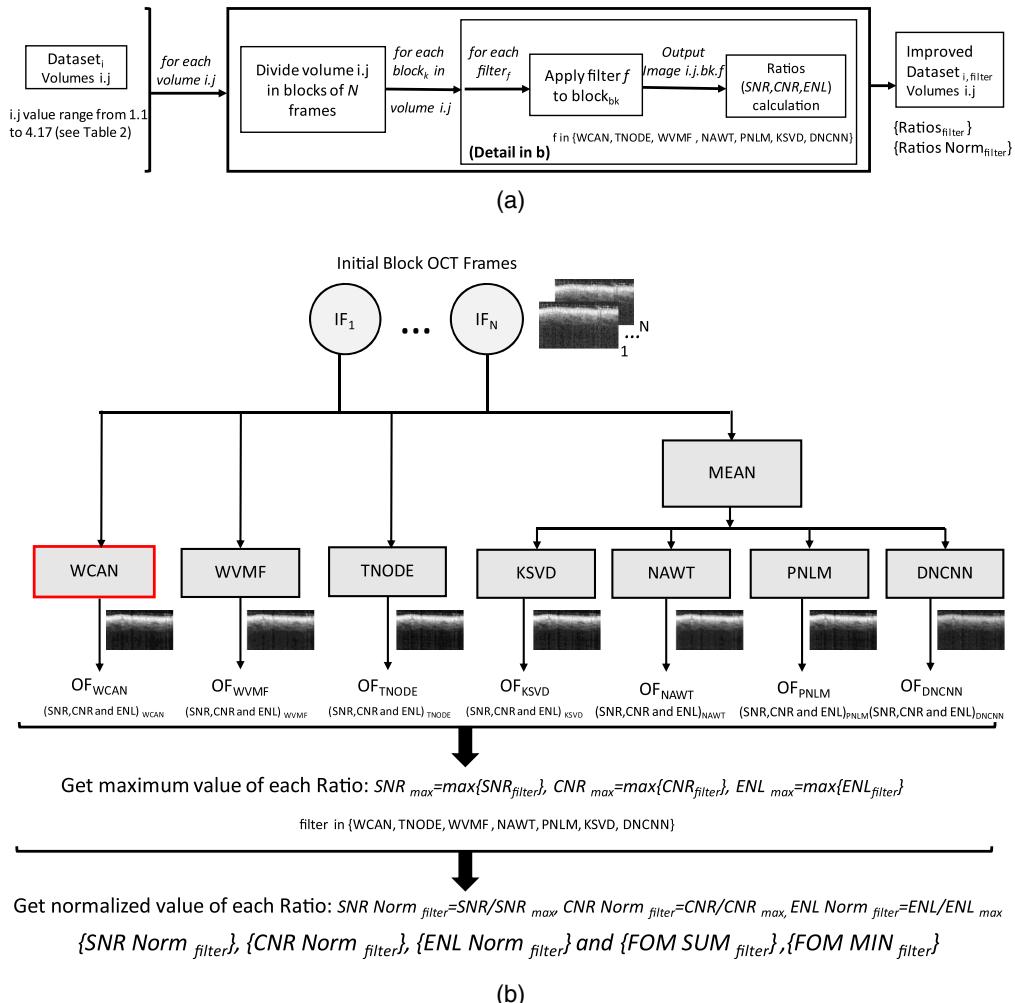
$$FOM_{MIN} = \text{Min}(SNR_{Norm}, CNR_{Norm}, ENL_{Norm}), \quad (15)$$

where *Norm* refers to the normalization of the metric, that is, the method that performed the best in, for example, the *SNR* criteria, would have an  $SNR_{Norm}$  equal to one. Therefore, an  $FOM_{SUM}$  of three would indicate that the method performed the best in all image-quality metrics (*SNR*, *CNR*, and *ENL*). The metric  $FOM_{MIN}$  is used to assess the robustness of the methods and to detect unbalanced combinations of *SNR*, *CNR*, and *ENL*.

We compared the performance of the WCAN algorithm with respect to other state-of-the-art methods via two different approaches. First, we used 2D filters that have previously demonstrated excellent performance in OCT speckle reduction such as PNLM,<sup>32</sup> complex wavelet based K-SVD (KSVD),<sup>46</sup> and noise adaptive wavelet thresholding (NAWT)<sup>43</sup> or in general image denoising (DNCNN).<sup>34</sup> Next, we used two 3D filters: i.e., the wavelet multiframe algorithm (WVMF)<sup>45</sup> and TNODE.<sup>47</sup> Table 4 gives sources for implementations of these methods. Figure 4 shows the process of evaluation for the various methods. All steps were performed over all frames in every dataset. The denoised OCT frames generated as outputs were compared using the *SNR*, *CNR*, and *ENL* performance metrics.

**Table 4** State-of-the-art denoising software used in the assessment of the proposed method.

Year	Denoising method	Official website
2012	WVMF <sup>45</sup>	Image denoising algorithms archive.
2015	KSVD <sup>46</sup>	State-of-the-art method for OCT denoising.
2016	PNLM <sup>32</sup>	PNLM: a probability-based non-local means filter for speckle noise suppression in optical coherence tomography images.
2017	DNCNN <sup>34</sup>	Beyond gaussian denoiser: residual learning of deep CNN for image denoising.
2017	NAWT <sup>43</sup>	Implemented following the authors' manuscript.
2018	TNODE <sup>47</sup>	Volumetric non-local-means based speckle reduction for optical coherence tomography.



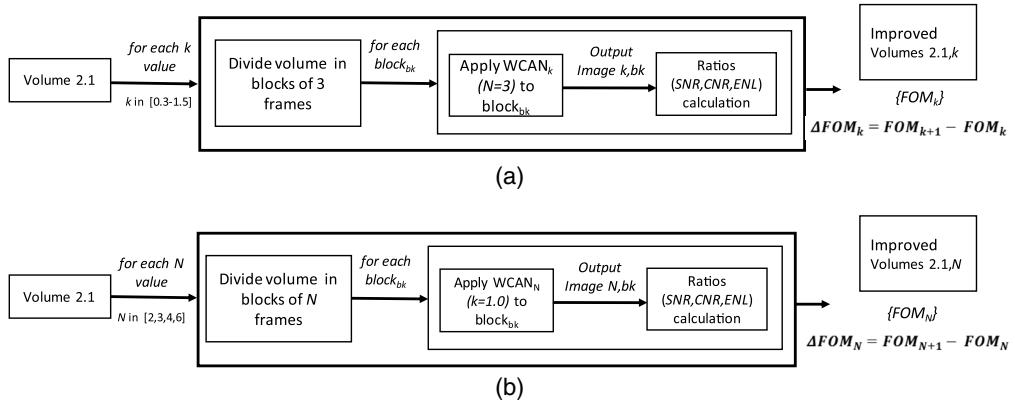
**Fig. 4** (a) Process used for the evaluation of all methods with all datasets. Each volume in the study was divided into blocks of  $N$  frames. For each block, all filters were applied and the evaluation metrics were computed. Finally, all blocks were combined and the improved version of each volume and each filter was created. (b) Details of the process for each block. The compounding of each block with  $N$  initial frames ( $\text{IF}_1, \dots, \text{IF}_N$ ) was performed using WCAN, WVMF, TNODE, and MEAN. The output of the MEAN was applied to the methods KSVD, NAWT, PNLM, and DNCNN. The output frames from each method ( $\text{OF}_{\text{WCAN}}, \text{OF}_{\text{WVMF}}, \text{OF}_{\text{TNODE}}, \text{OF}_{\text{KSVD}}, \text{OF}_{\text{NAWT}}, \text{OF}_{\text{PNLM}}$ , and  $\text{OF}_{\text{DNCNN}}$ ), as applied to all frames, were used in the quantitative assessment. The normalized value of the metrics (*SNR Norm*, *CNR Norm*, *ENL Norm*, and *FOM*) were calculated with respect to the maximum value of each metric from the output frames. We used these normalized values for all volumes in comparing the quantitative assessments for all filters.

### **3 Results and Discussion**

### 3.1 Parameter Assessment

Before applying the method to all datasets and comparing it with the state-of-the-art filters, we evaluated the impact of the two main parameters of the algorithm, i.e.,  $k$  [the threshold parameter in Eq. (3)] and  $N$  (the number of frames to consider in the compounding process). We used Volume 2.1 (see Table 2) and the reference frames from Volume R.2 (Table 3) for this purpose in the two experiments outlined in Fig. 5.

First, we varied the parameter  $k$  from 0.3 to 1.5 in steps of 0.1, with  $N$  set to 3. In the second experiment, we varied the parameter  $N$  from 2 to 6, with  $k$  set to 1.0. After applying the WCAN method in both cases, we calculated the quantitative metrics (*SNR*, *CNR*, and *ENL*) and the resulting *FOM<sub>SUM</sub>* (only *FOM* for the rest of this section). We then computed the difference in the *FOM*



**Fig. 5** Process followed for the assessment of the parameters  $k$  and  $N$  using Volume 2.1 from the MUW-2 dataset and the reference frames from Volume R.2. (a) Evaluation of the parameter  $k$  with  $N = 3$  for values of  $k$  from 0.3 to 1.4 in steps of 0.1. We divided the volume into blocks of three frames and applied WCAN for each  $k$  value. After calculating the quantitative metrics, we compared the resulting  $FOM$ s of the enhanced frames for consecutive values of  $k$ . (b) Evaluation of the parameter  $N$  with  $k = 1$  for values of  $N$  from 2 to 6. For each  $N$  value, we divided the volume into blocks of  $N$  frames and applied WCAN with  $k$  set to 1. After calculating the quantitative metrics, we compared the resulting  $FOM$  of the enhanced frames for consecutive values of  $N$ .

values for each enhanced frame for two consecutive values of  $k$  (such as  $k = 0.4$  and 0.3) and  $N$  (such as  $N = 3$  and  $N = 2$ ). This process generated a set of  $FOM$ -value differences  $\Delta FOM_K$  for each pair of consecutive  $k$  values and a set of  $\Delta FOM_N$  for each pair of consecutive  $N$  values.

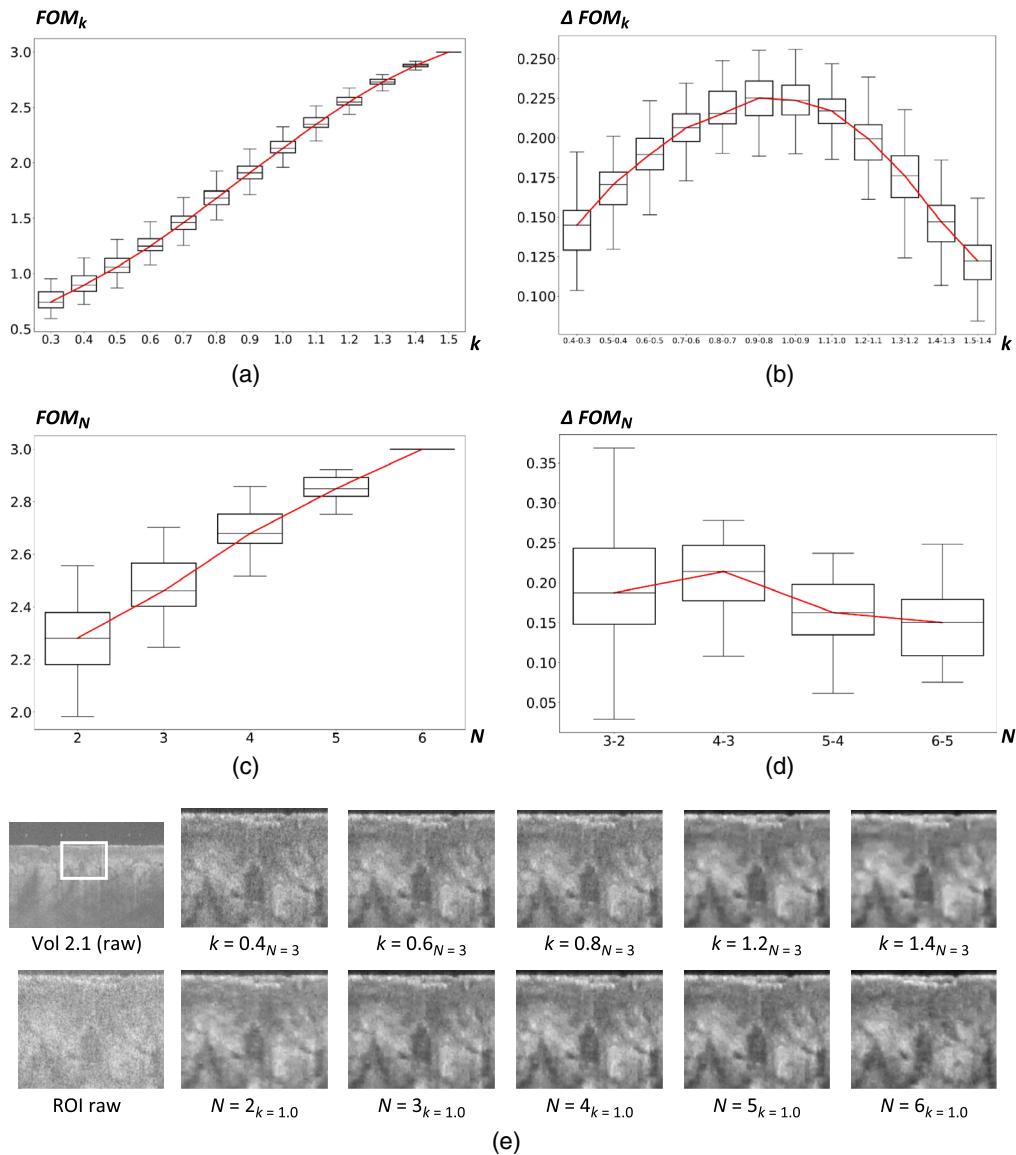
Figure 6 gives the main results and some examples of the application of the WCAN method to Volume 2.1. We observe that the improvement in the quantitative metrics reaches a maximum between  $k = 0.8$  and 0.9, with a mean improvement in  $FOM$  of 0.213, and any further increase in  $k$  decreases the incremental improvement. Note that this does not imply that  $k = 0.9$  is the optimal value to use but that  $k = 0.9$  is the value for which we have the greatest improvement with respect to the previous value of  $k$ . Although increasing the value of  $k$  always produces a quantitative improvement in the image, there are diminishing returns from increasing  $k$  beyond  $k = 0.9$ . We can observe the same trend for the parameter  $N$ , the maximum mean improvement of  $FOM$ , i.e.,  $\Delta FOM_N = 0.229$ , occurs between  $N = 4$  and  $N = 3$ . The image examples in Fig. 6 show how increases in  $N$  and  $k$  produce better noise reduction in the resultant image. One of the main advantages of the proposed method is that it is easily adaptable to very different OCT settings by simply adjusting these two parameters.

To find appropriate values of  $k$  and  $N$  for the quantitative assessment, we used the first volume in each dataset (Volume IDs 1.1, 2.1, 3.1, and 4.1 for the MUW-1, MUW-2, MUW-3, and A2ASDOCT datasets, respectively, in Table 2 and the references frames with Volumes IDs R.1, R.2, R.3, and R.4 I in Table 3). Table 5 gives the final values chosen for each dataset, having been determined empirically as a suitable balance between performance metric improvement, qualitative visual quality, and execution time.

We also evaluated the influence of the wavelet family in the calculation of 2D discrete stationary wavelet transform. We tested the wavelets families Haar, Daubechies (db1-db10), Symlets (sym2-sym8), Coiflets (coif1-coif5), and BiorSplines (bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, and bior6.8) using the Volume ID 2.1 and the reference set R.2 of Table 3. The Table 6 presents the results of the wavelet with the best performance in each family. As we can observe, the Haar wavelet shows the best performance and was the wavelet selected for the rest of the experiments.

### 3.2 Quantitative Evaluation

We now present the results of our quantitative evaluation of the application of the compared methods to the complete stack of frames for all four datasets in this study. In the evaluation, we used the same ROIs for all frames in the same volume to maintain a consistent reference (Fig. 7 shows one frame with the ROIs used for Volume 1.1). The detailed values of the metrics



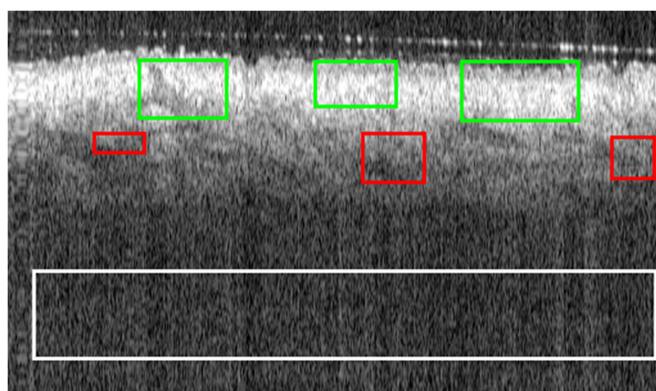
**Fig. 6** Results of the evaluation of the parameters  $k$  and  $N$  for the WCAN method, using the Volume 2.1 from the MUW-2 dataset and the reference frames from Volume R.2. (a)  $FOM$  for consecutive  $k$  values from 0.3 to 1.5. (b) Improvement in the  $FOM$  metric for consecutive  $k$  values from 0.3 to 1.5. The horizontal axis represents consecutive pairs of values for  $k$ . The vertical axis represents the  $FOM$  differences over all frames in the volume for consecutive pairs of values for  $k$ . (c)  $FOM$  for consecutive  $N$  values from 2 to 6. (d) Improvement in  $FOM$  metric for consecutive  $N$  values from  $N = 2$  to  $N = 6$ . The horizontal axis represents consecutive pairs of values for  $N$ . The vertical axis represents the  $FOM$  differences over all frames in the volume for consecutive pairs of values for  $N$ . (e) Examples of the application of the WCAN method for different combinations of  $k$  and  $N$  in the ROI marked as a white box in the top-left raw noisy OCT image.

**Table 5** Parameters used in the assessment of the WCAN method for each dataset.  $N$  is the number of frames considered at consecutive positions in the compounding process.  $k$  is the adjustment parameter for the thresholding process. The number of decomposition levels,  $L$ , was set to four for all datasets.

Parameter	MUW-1 dataset	MUW-2 dataset	MUW-3 dataset	A2ASDOCT dataset
$N$	3	3	2	4
$k$	1	1	1.1	1.6

**Table 6** Quantitative evaluation of the wavelet families Haar, Daubenchies, Symlets, Coiflets and BiorSplines, considering the Volume ID 2.1 and the reference set R.2 from Table 3. The results show the mean  $\pm$  standard deviation of the improvement with respect to the initial metrics presented in Table S1 in the [Supplementary Material](#) (raw images).

Wavelet	SNR	CNR	ENL
haar	$12.30 \pm 2.44$	$2.02 \pm 0.36$	$622.80 \pm 101.53$
db2	$11.78 \pm 2.25$	$1.86 \pm 0.31$	$537.78 \pm 82.06$
sym3	$11.07 \pm 2.03$	$1.85 \pm 0.32$	$571.68 \pm 79.25$
coef1	$11.74 \pm 2.23$	$1.85 \pm 0.31$	$533.41 \pm 81.91$
bior3.3	$11.71 \pm 2.21$	$1.82 \pm 0.30$	$521.72 \pm 81.37$



**Fig. 7** OCT raw image and ROIs used for the calculation of the quality metrics. The white rectangular region was used for the noise estimation; the red rectangles represent the homogeneous regions ( $H = 3$ ) used to calculate  $ENL$ , and the green rectangles represent the nonhomogeneous regions. The sum of both individual metrics is used to calculate the  $CNR$  ( $R = 6$ ).

**Table 7** Quantitative evaluation for the MUW-1 dataset, considering the results of the application of all filters to Volumes 1.1–1.5 (see Table 2) and the reference frames R.1 (see Table 3). The compounding was performed using three B-scans per position.

Filter	SNR Norm	CNR Norm	ENL Norm	FOM <sub>SUM</sub>	FOM <sub>MIN</sub>
WCAN	<b><math>0.99 \pm 0.026</math></b>	<b><math>0.97 \pm 0.065</math></b>	$0.91 \pm 0.117$	<b><math>2.87 \pm 0.192</math></b>	<b><math>0.91 \pm 0.117</math></b>
TNODE	$0.88 \pm 0.044$	$0.93 \pm 0.057$	<b><math>0.95 \pm 0.101</math></b>	$2.77 \pm 0.148$	$0.86 \pm 0.075$
PNLM	$0.75 \pm 0.102$	$0.80 \pm 0.067$	$0.61 \pm 0.168$	$2.16 \pm 0.271$	$0.61 \pm 0.165$
NAWT	$0.87 \pm 0.093$	$0.89 \pm 0.083$	$0.82 \pm 0.126$	$2.58 \pm 0.262$	$0.80 \pm 0.119$
WVMF	$0.82 \pm 0.051$	$0.78 \pm 0.064$	$0.68 \pm 0.101$	$2.28 \pm 0.166$	$0.68 \pm 0.101$
KSVD	$0.80 \pm 0.085$	$0.73 \pm 0.064$	$0.64 \pm 0.105$	$2.18 \pm 0.169$	$0.64 \pm 0.101$
DNCNN	$0.50 \pm 0.051$	$0.54 \pm 0.059$	$0.44 \pm 0.077$	$1.48 \pm 0.137$	$0.42 \pm 0.066$

for all datasets are included in Tables S2–S7 in the [Supplementary Material](#). In Tables 7–10 and Fig. 8, we present the aggregate results. Tables 7–10 show the mean  $\pm$  standard deviation of the improvement with respect to the initial metrics presented in Table S1 in the [Supplementary Material](#) (raw images). The values corresponding to the best performance for the individual metrics and for the  $FOMs$  are shown in bold.

**Table 8** Quantitative evaluation for the MUW-2 dataset, considering the results of the application of all filters to Volumes 2.1–2.5 (see Table 2) and the reference frames R.2 (see Table 3). The compounding was performed using three B-scans per position.

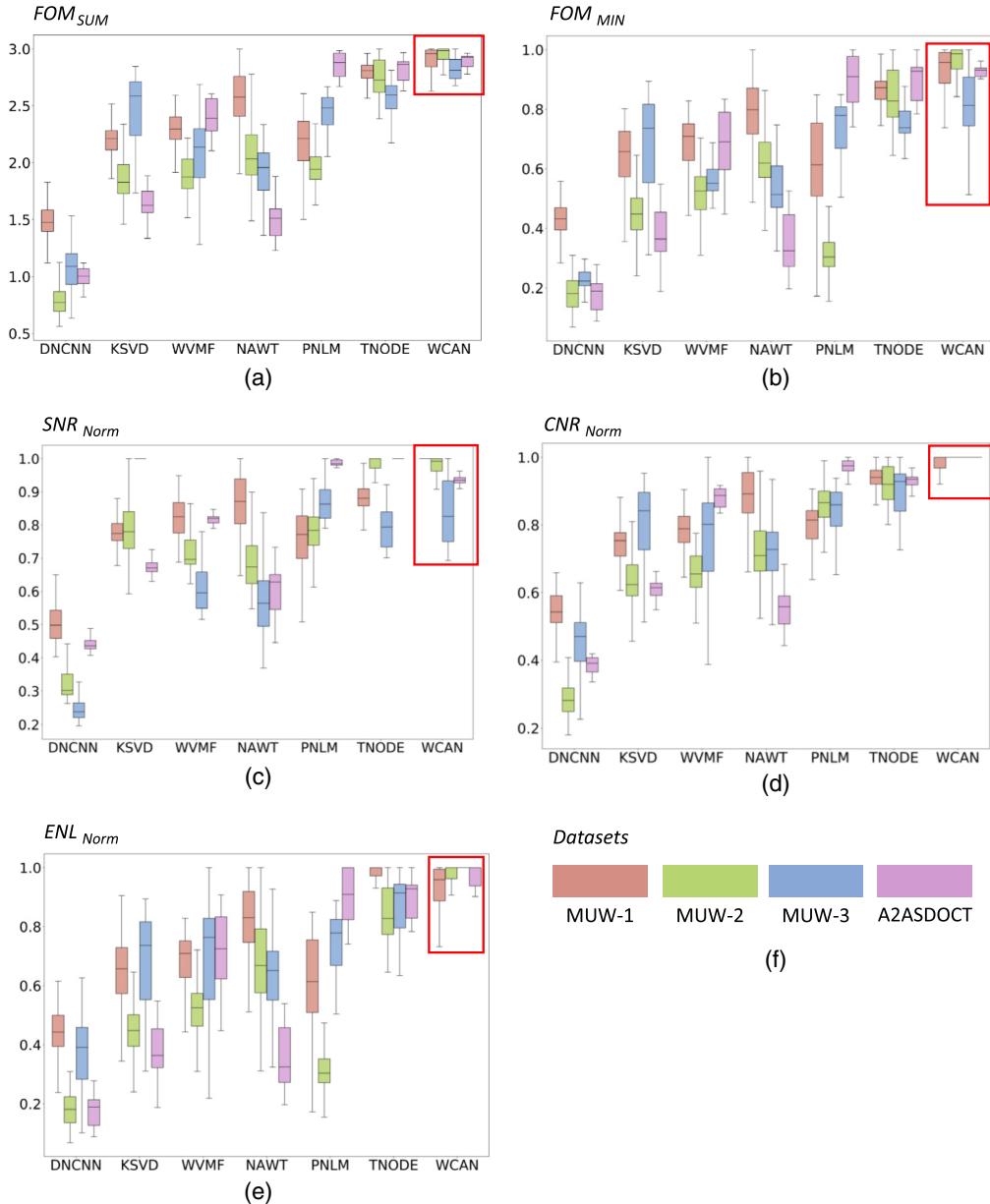
Filter	<i>SNR Norm</i>	<i>CNR Norm</i>	<i>ENL Norm</i>	<i>FOM<sub>SUM</sub></i>	<i>FOM<sub>MIN</sub></i>
WCAN	0.97 ± 0.064	<b>0.98 ± 0.055</b>	<b>0.96 ± 0.075</b>	<b>2.91 ± 0.187</b>	<b>0.95 ± 0.076</b>
TNODE	<b>0.98 ± 0.026</b>	0.92 ± 0.055	0.85 ± 0.100	2.75 ± 0.165	0.85 ± 0.099
PNLM	0.78 ± 0.075	0.87 ± 0.055	0.31 ± 0.062	1.96 ± 0.142	0.31 ± 0.062
NAWT	0.68 ± 0.076	0.74 ± 0.116	0.69 ± 0.178	2.11 ± 0.335	0.62 ± 0.114
WVMF	0.72 ± 0.053	0.66 ± 0.057	0.52 ± 0.086	1.89 ± 0.153	0.52 ± 0.086
KSVD	0.78 ± 0.083	0.64 ± 0.080	0.45 ± 0.085	1.87 ± 0.213	0.45 ± 0.085
DNCNN	0.32 ± 0.048	0.29 ± 0.048	0.18 ± 0.052	0.79 ± 0.121	0.18 ± 0.052

**Table 9** Quantitative evaluation for the MUW-3 dataset, considering the results of the application of all filters to Volumes 3.1–3.8 (see Table 2) and the reference frames R.3 (see Table 3). The compounding was performed using two B-scans per position.

Filter	<i>SNR Norm</i>	<i>CNR Norm</i>	<i>ENL Norm</i>	<i>FOM<sub>SUM</sub></i>	<i>FOM<sub>MIN</sub></i>
WCAN	0.85 ± 0.102	<b>1.00 ± 0.027</b>	<b>0.99 ± 0.040</b>	<b>2.84 ± 0.112</b>	<b>0.84 ± 0.097</b>
TNODE	0.80 ± 0.063	0.90 ± 0.070	0.87 ± 0.103	2.56 ± 0.144	0.77 ± 0.045
PNLM	0.87 ± 0.055	0.84 ± 0.071	0.74 ± 0.104	2.46 ± 0.150	0.31 ± 0.052
NAWT	0.58 ± 0.116	0.72 ± 0.095	0.64 ± 0.128	1.95 ± 0.267	0.55 ± 0.092
WVMF	0.61 ± 0.066	0.75 ± 0.139	0.68 ± 0.194	2.05 ± 0.313	0.51 ± 0.068
KSVD	<b>0.98 ± 0.049</b>	0.81 ± 0.112	0.69 ± 0.156	2.48 ± 0.302	0.45 ± 0.077
DNCNN	0.25 ± 0.040	0.44 ± 0.115	0.37 ± 0.138	1.07 ± 0.254	0.18 ± 0.041

**Table 10** Quantitative evaluation for the A2ASDOCT dataset, considering the results of the application of all filters to the 17 retinal sets (see Table 2) and the reference frame R.4 (see Table 3). The compounding was performed using four B-scans per position.

Filter	<i>SNR Norm</i>	<i>CNR Norm</i>	<i>ENL Norm</i>	<i>FOM<sub>SUM</sub></i>	<i>FOM<sub>MIN</sub></i>
WCAN	0.93 ± 0.040	<b>1.00 ± 0.009</b>	<b>0.96 ± 0.065</b>	<b>2.89 ± 0.073</b>	<b>0.91 ± 0.057</b>
TNODE	0.99 ± 0.028	0.93 ± 0.024	0.90 ± 0.067	2.82 ± 0.098	0.89 ± 0.061
PNLM	<b>0.99 ± 0.009</b>	0.97 ± 0.026	0.90 ± 0.091	2.86 ± 0.112	0.89 ± 0.083
NAWT	0.60 ± 0.074	0.55 ± 0.060	0.35 ± 0.107	1.50 ± 0.171	0.35 ± 0.099
WVMF	0.81 ± 0.055	0.88 ± 0.027	0.70 ± 0.139	2.39 ± 0.155	0.67 ± 0.122
KSVD	0.67 ± 0.046	0.60 ± 0.037	0.38 ± 0.096	1.65 ± 0.135	0.38 ± 0.096
DNCNN	0.43 ± 0.044	0.38 ± 0.026	0.18 ± 0.058	1.00 ± 0.082	0.18 ± 0.059

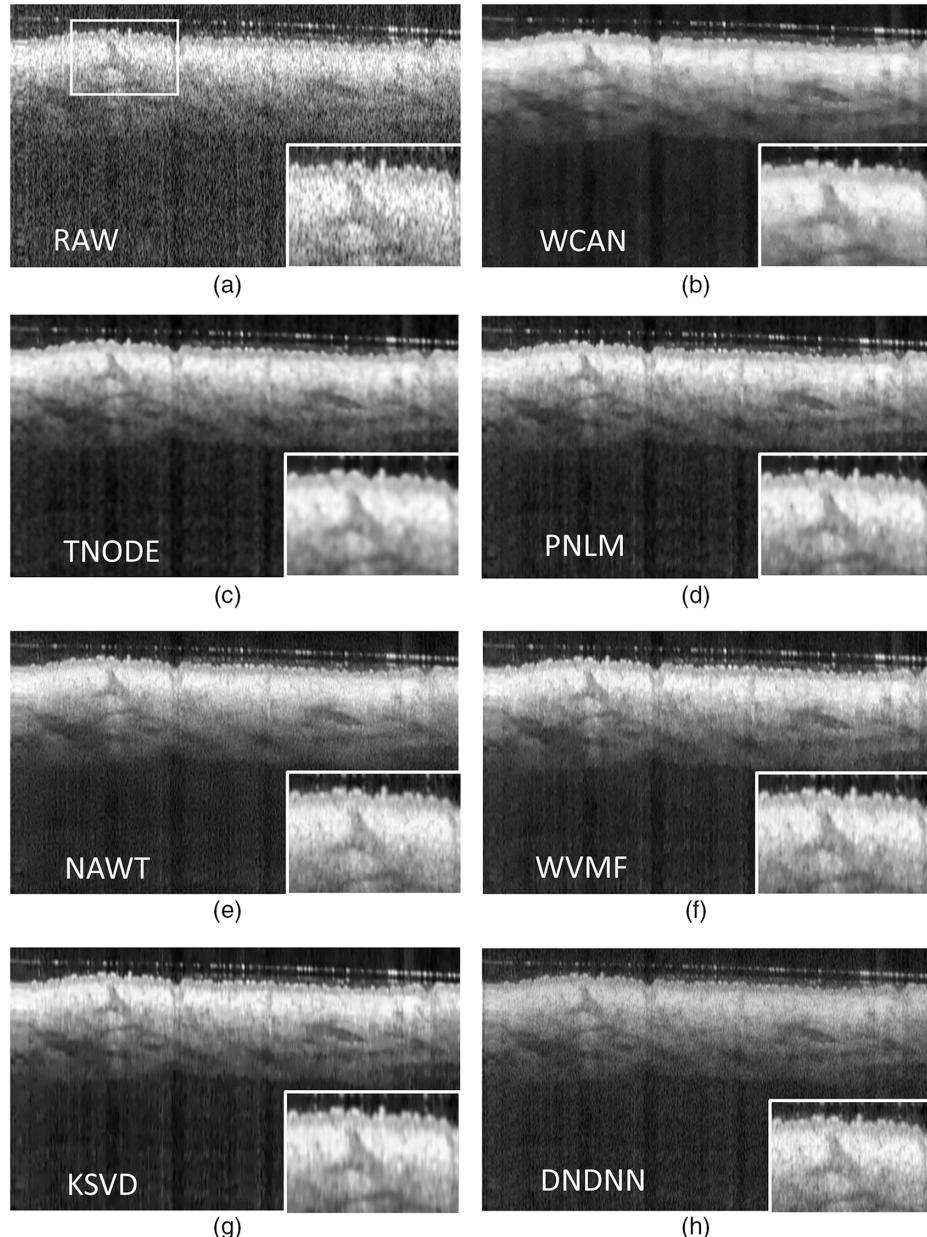


**Fig. 8** Value distribution for the different metrics of the quantitative evaluation of all datasets used in the study (see Table 2). The proposed WCAN method is shown in the rightmost column within the red box. (a)  $FOM_{SUM}$  calculated using Eq. (14). (b)  $FOM_{MIN}$  calculated using Eq. (15). (c)  $SNR\ Norm$  calculated using Eq. (11) and normalized with respect to the maximum value of the metric for each denoised frame. (d)  $CNR\ Norm$  calculated using Eq. (12) and normalized with respect to the maximum value of the metric for each denoised frame. (e)  $ENL\ Norm$  calculated using Eq. (13) and normalized with respect to the maximum value of the metric for each denoised frame. (f) Order of the datasets for each method.

To facilitate comparison between metrics, we used the normalized value of each metric, calculated from the maximum values for all filtered images in each group of frames [see Fig. 4(b)]. We can then refer to metric values corresponding to the best filtered output for each group of frames to make fair comparisons across the volume. We should first note that the best results across all datasets were accomplished by WCAN, with the highest values for  $FOM_{SUM}$  and standard deviations between the smallest (MUW-3) and the second smallest (MUW-1, MUW-2, and A2ASDOCT). We have the same situation with the metric  $FOM_{MIN}$ , where WCAN presents the highest values. The TNODE algorithm had the second-best performance

across all datasets. WCAN was also the best option, in most cases, when considering each metric separately. It returned the best *CNR* and *ENL* results for all datasets (except for MUW-1, where it had the second highest *ENL*). WCAN had the best *SNR* score for MUW-1, was second for MUW-2, and third for MUW-3 and A2ASDOCT.

Another notable aspect of the WCAN results is its consistent performance with very different OCT image settings, image samples, and locations on the skin (see Table 2) and retina. As is clear from Figs. 8(a) and 8(b), other filters such as PNLM and KSVD perform very differently for the retina dataset (A2ASDOCT) than for the skin datasets (MUW-1, MUW-2, and MUW-3).

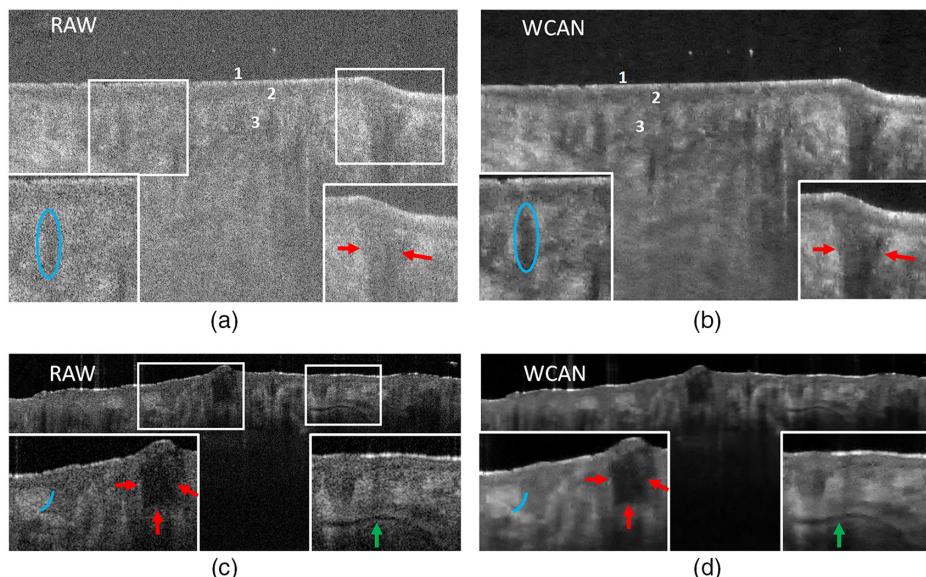


**Fig. 9** Comparison of the application of speckle reduction methods in a frame with  $1100 \times 512$  pixels from Volume 1.1 using three B-scans at consecutive positions and the reference frames R.1. (a) Raw OCT image with a detailed ROI. The remainder of the image is the filtered OCT image after applying the method. (b) Proposed WCAN method. (c) TNODE method. (d) PNLM method. (e) NAWT method. (f) WVMF method. (g) KSVD method. (h) DNCNN method.

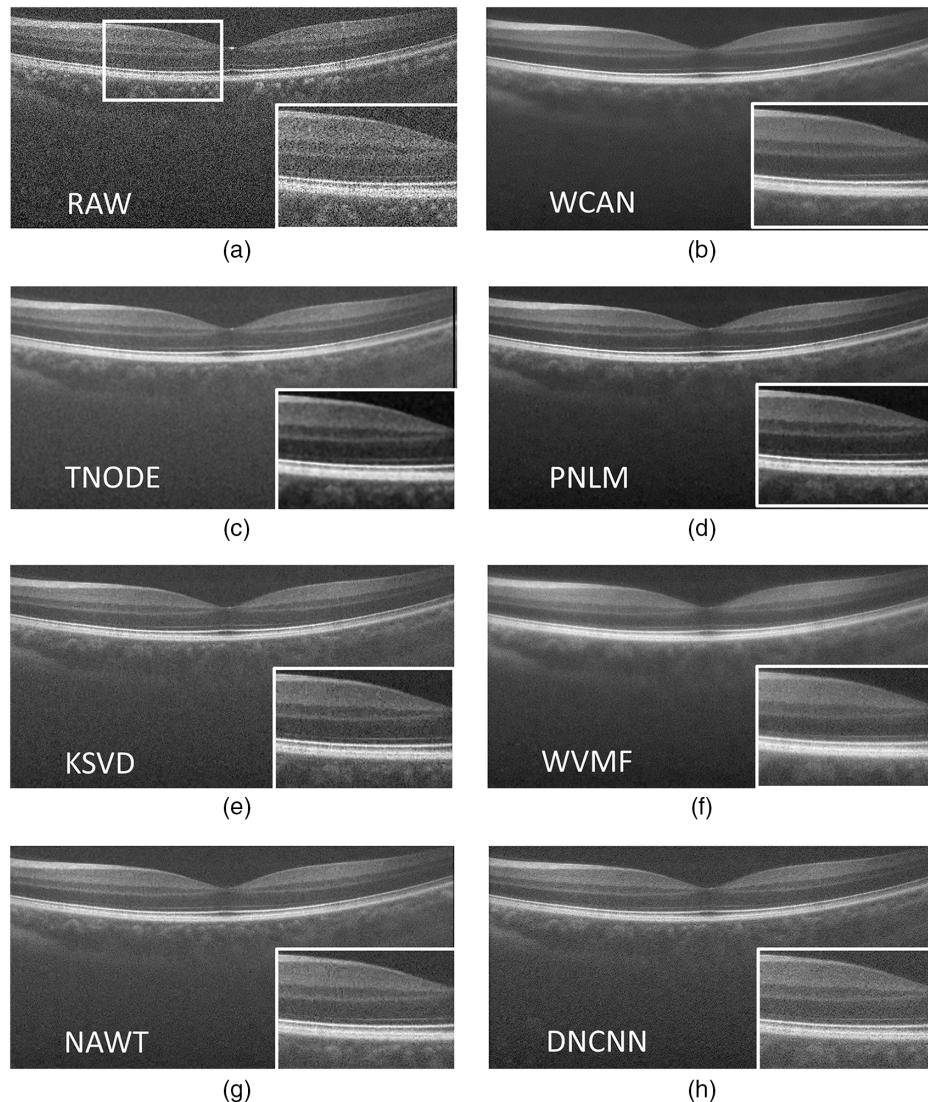
Figure 9 shows the significant differences in the OCT images generated by applying the state-of-the-art methods to Volume 1.1 and the reference frames R.1. Videos S1–S4 (see Figs. 13–16 in the Appendix) show the results of the application of the method over frames of the datasets MUW-1, MUW-2, MUW-3, and A2SDOCT.

Simple averaging was also evaluated. We used three volumes from datasets MUW-1, MUW-2, and MUW-3, where we have enough frames to perform the experiment in each volume. In particular, we considered Volumes ID 1.1, 2.1, 3.1 (Table 2) and the reference sets R.1, R.2, and R.3 of Table 3, respectively. We applied the WCAN method to each volume using the parameters previously identified in Table 5. In particular, we applied the WCAN to the three first frames of Volume ID 1.1 and 2.1 with  $k = 1$  and to the first two frames of Volume ID 3.1 with  $k = 1.1$ . Then, we performed the average of  $N$  frames iteratively until the metric  $FOM_{SUM}$  of averaging frames improved the metric of WCAN previously calculated. The results showed that to improve WCAN we needed to average 12, 27, and 13 frames for the Volumes ID 1.1, 2.1, and 3.1, respectively.

Figure 10 shows an analysis of some examples of the WCAN algorithm's performance with skin images corresponding to frames of the MUW-2 and MUW-3 datasets. The first sample (Volume 2.4) provides information about a superficial BCC on the head. In Figs. 10(a) and 10(b), speckle reduction using WCAN [Fig. 10(b)] enables a stronger identification of the skin layers (marked with digits 1–3) than does the raw image [Fig. 10(a)]. In addition, the stratum corneum layer and the epidermis–dermis interface are more clearly distinguished. The boundaries and contrast are also enhanced, shown by comparing the area enclosed by the blue ellipse and the red arrows in the zoom-in boxes in the raw image [Fig. 10(a)] and after filtering [Fig. 10(b)]. The second example (Volume 3.8) shows a cheek with folliculitis. Figures 10(c) and 10(d) show a sharper difference between the regions indicated by the blue line and the red arrows and the structure indicated by the green arrow in the zoom-in boxes.



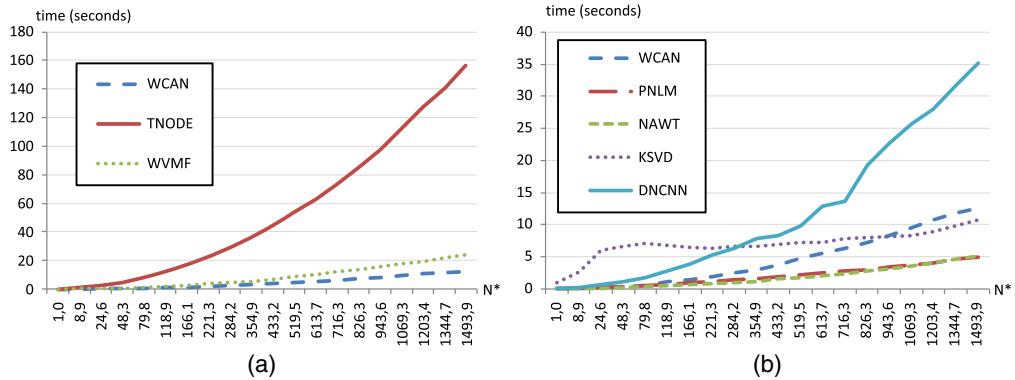
**Fig. 10** Comparison of the application of the WCAN method to frames from the MUW-2 and MUW-3 datasets. (a) Raw OCT skin image from a  $1000 \times 680$  pixel sample of a head affected by BCC (Volume 2.4). Two detail ROIs are identified by white rectangular boxes. The digits 1, 2, and 3 in the image identify the stratum corneum, epidermis, and dermis layer, respectively. The blue ellipse and red arrows indicate regions that show the contrast enhancement obtained by filtering the raw image. (b) WCAN image processed using three consecutive frames from Volume 2.4, starting from the a frame. (c) Raw OCT skin image from a  $1038 \times 460$  pixel sample of a cheek with folliculitis (Volume 3.8). Two zoom-in ROIs are identified by white rectangular boxes. The blue line and the red and green arrows identify different areas affected by speckle reduction after application of the proposed method. (d) WCAN image processed using two consecutive frames from Volume 3.8 starting from the c frame.



**Fig. 11** Comparison of the application of speckle reduction methods in a  $900 \times 450$ -pixel frame from Volume 4.1 using four B-scans at consecutive positions. (a) Raw OCT image and a detail ROI used for the qualitative comparison of all methods. (b) Proposed WCAN method. (c) TNODE method. (d) PNLM method. (e) KSVD method. (f) WVMF method.

Figure 11 enables qualitative comparisons for a retinal image from the A2ASDOCT database after applying each of the denoising methods in the study. Figure 11 shows that the WCAN algorithm gives significant noise suppression, which enables clear identification of the boundaries between retinal layers.

Finally, to assess the computation time for the various algorithms we applied all methods of the study to 20 scaled versions of the first two frames of Volume ID 1.1. For the WCAN method we used  $k = 1$  and the reference frames R.1 from Table 3. The original size of the volume was two frames with  $1100 \times 512$  pixels. We used scales from 0.1 to 4 in steps of 0.2 (20 scales in total) to resize the frames. For each scale we applied all of the methods and calculated the execution time using the method `timeit` of Matlab. In Fig. 12 we present the results. As shown in Fig. 12(a) the application of the WCAN method requires less execution time than the other pure compounding alternatives such as TNODE and WVMF. The PNLM and NAWT methods presented the best performance among all of the algorithms tested [Fig. 12(b)]. Finally, to estimate the computational complexity of the algorithm, we applied a basic fitting of the performance metrics of WCAN, resulting a linear polynomial function of the type  $y = p_1 * x + p_2$  with  $p_1 = 0.0087$ ,  $p_2 = 0.0086$ , and the norm of residuals equaling 0.699. So we can estimate



**Fig. 12** Comparison of the computation time in seconds of all methods of the study (WCAN, TNODE, WVMF, PNLM, NAWT, KSVD, and DNCNN). The parameter  $N^*$  represents the total number of the pixels considered in each test ( $1100 \times 512 \times 2 \times$  scale) normalized with respect to the smallest one ( $1100 \times 512 \times 2 \times 0.1$ ). (a) Computation time comparison of WCAN with respect to the compounding methods TNODE and WVMF. (b) Computation time comparison of WCAN with respect to PNLM, NAWT, KSVD, and DNCNN methods.

that the time complexity of the method is linear. The tests were executed in Matlab 2018b (MathWorks, Inc.) on a personal computer (Intel 3.3 GHz CPU, 32 GB memory).

## 4 Conclusion

This paper has described a speckle reduction method for OCT volumes. The method is based on a multiscale adaptive noise-wavelet-compounding strategy. The method requires the estimation of the noise variance in the wavelet domain of the OCT setting. In this study, we used previously acquired images of homogeneous scattering examples. The noise variance is used to compute the weights for the detail coefficients of each subband of each frame in the volume before compounding all of the filtered frames.

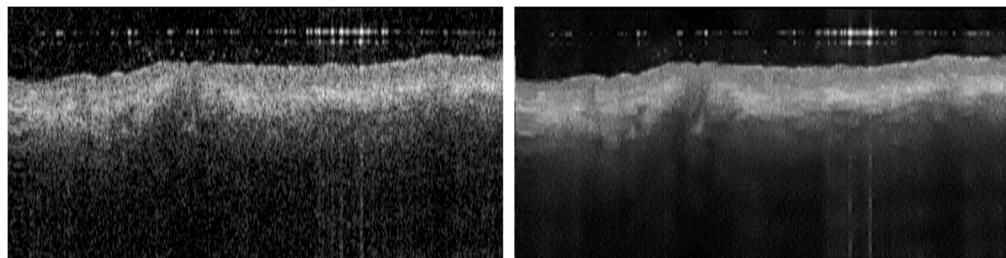
One of the major highlights of our method, and a novel contribution in the field of wavelet denoising, is the introduction of a new compounding scheme that integrates the noise variance influence of the input frames in relation to the variance of a set of reference frames. With this new approach, a visual inspection of the denoised frames presented shows the potential of the proposed algorithm to reduce the speckle noise efficiently without generating relevant denoising artifacts or producing a degradation of the structures present in the image. The simplicity, the ability of the method to adapt to different environments (OCT settings and noise levels), the robustness above all metrics considered, and its reduced computational time are its most significant advantages. Another potential benefit is the capacity of the method to adjust the amount of noise reduction by the modification of the parameter  $k$ , which may facilitate the adaptation to different clinical needs. The main weaknesses of the method is the need for a prior acquisition of a homogeneous sample to determine the noise variance used in the adaptive compounding of the detail coefficients in the wavelet decomposition. Nevertheless, as described in the estimation of the noise in datasets MUW-2 and A2ASDOCT, this drawback can be avoided through the use of a homogeneous ROI of the input frames. The other line of improvement is the automatic assessment of the parameters of the method ( $k$  and the number of frames of the input and reference frames). However, this assessment is needed only once when fine tuning the algorithm in a particular OCT setting.

A thorough evaluation was performed by applying the method to four different OCT settings, i.e., around 18 skin volumes in three datasets and 17 retinal volumes in a fourth dataset. We compared our method's results with those from six state-of-the-art algorithms. The results of a quantitative evaluation based on five different metrics demonstrated that the proposed method achieved the best performance among the tested techniques in suppressing noise. A qualitative visual comparison of images confirmed the relative performance of the proposed method and suggests the potential application of the proposed OCT denoising technique to improving image

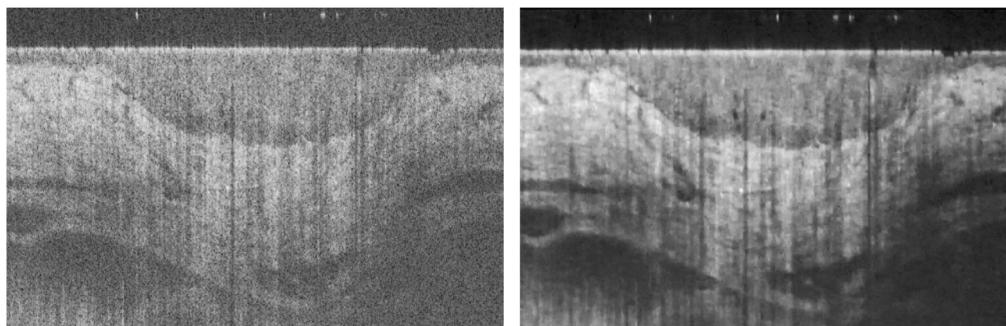
quality prior to clinical diagnosis based on visual assessment. Future work will involve an automatic process for the assessment of the parameters of the method ( $k$  and the number of frames of the input and reference frames) through the training of a machine learning model based on quantitative metrics. Furthermore, a formal qualitative evaluation will be performed by clinical experts tackling the specific needs in their clinical routine, such as the capacity of the method to preserve or enhance the identification of tissue layers and their borders and the ability to visualize structures or to identify vasculatures.

## 5 Appendix

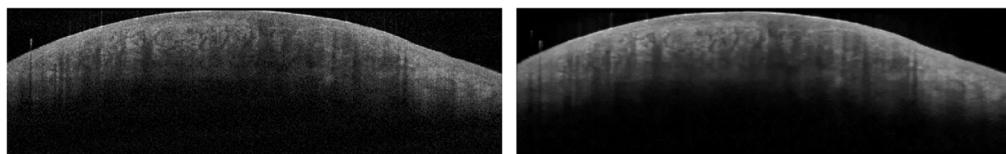
Videos S1–S4 (see Figs. 13–16) show the results of the application of the method over frames of the datasets MUW-1, MUW-2, MUW-3, and A2SDOCT.



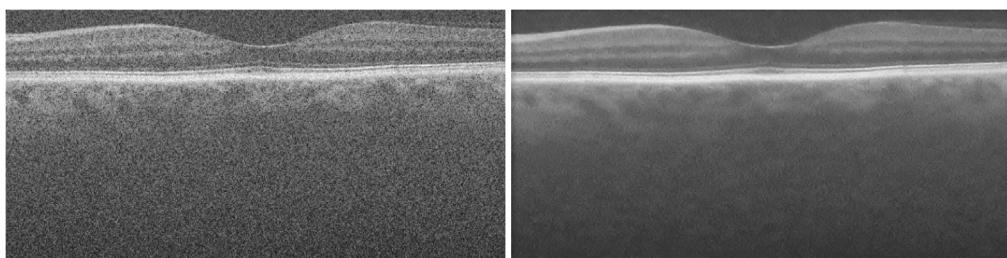
**Fig. 13** Results of the application of the WCAN method over frames from dataset MUW-1, Volumes 1.1–1.5 (see Table 2), using three B-scans at consecutive positions. (a) OCT raw image Volume 1.2 (b) WCAN image. (Video S1, MP4, 8.3MB [URL: <https://doi.org/10.1117/1.JBO.26.6.065001.1>]).



**Fig. 14** Results of the application of the WCAN method over frames from dataset MUW-2, Volumes 2.1–2.5 (see Table 2), using three B-scans at consecutive positions. (a) OCT raw image Volume 2.5 (b) WCAN image. (Video S2, MP4, 10.6 MB [URL: <https://doi.org/10.1117/1.JBO.26.6.065001.2>]).



**Fig. 15** Results of the application of the WCAN method over frames from dataset MUW-3, Volumes 3.1–3.5 (see Table 2), using two B-scans at consecutive positions. (a) OCT raw image Volume 3.3 (b) WCAN image. (Video S3, MP4, 1.8 MB [URL: <https://doi.org/10.1117/1.JBO.26.6.065001.3>]).



**Fig. 16** Results of the application of the WCAN method over frames from dataset A2ASDOCT, Volumes 4.1–4.17 (see Table 2), using four B-scans at consecutive positions. (a) OCT raw image Volume 4.12 (b) WCAN image. (Video S4, MP4, 9.9 MB [URL: <https://doi.org/10.1117/1.JBO.26.6.065001.4>]).

## Disclosures

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

## Acknowledgments

This work was partly supported by the European Union FP7 project BiopsyPen (Grant No. 611132), the European Regional Development Fund, and the Spanish Ministry of Science, Innovation and Universities (Grant Nos. TEC2015-66978-R and RTI2018-098682-B-I00 MCIU/AEI/FEDER, EU).

## Code, Data, and Materials Availability

We have made a Matlab implementation of the WCAN method available through Code Ocean, including the full code for the noise estimation and the denoising algorithm, together with raw volume data for testing purposes.

## References

1. D. Huang et al., “Optical coherence tomography,” *Science* **254**(5035), 1178–1181 (1991).
2. G. J. Jaffe and J. Caprioli, “Optical coherence tomography to detect and manage retinal disease and glaucoma,” *Am. J. Ophthalmol.* **137**(1), 156–169 (2004).
3. H. G. Bezerra et al., “Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications,” *JACC: Cardiovasc. Interventions* **2**(11), 1035–1046 (2009).
4. T. Gambichler, A. Pljakic, and L. Schmitz, “Recent advances in clinical application of optical coherence tomography of human skin,” *Clin. Cosmet. Investig. Dermatol.* **8**, 345–354 (2015).
5. J. Olsen, J. Holmes, and G. B. E. Jemec, “Advances in optical coherence tomography in dermatology—a review,” *J. Biomed. Opt.* **23**(4), 040901 (2018).
6. A. A. Hussain, L. Themstrup, and G. B. E. Jemec, “Optical coherence tomography in the diagnosis of basal cell carcinoma,” *Arch. Dermatol. Res.* **307**(1), 1–10 (2015).
7. S. A. Alawi et al., “Optical coherence tomography for presurgical margin assessment of non-melanoma skin cancer—a practical approach,” *Exp. Dermatol.* **22**(8), 547–551 (2013).
8. O. Markowitz et al., “Evaluation of optical coherence tomography as a means of identifying earlier stage basal cell carcinomas while reducing the use of diagnostic biopsy,” *J. Clin. Aesthet. Dermatol.* **8**(10), 14–20 (2015).
9. J. M. Schmitt, S. H. Xiang, and K. M. Yung, “Speckle in optical coherence tomography,” *J. Biomed. Opt.* **4**(1), 95–105 (1999).

10. M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," *Opt. Lett.* **25**(8), 545–547 (2000).
11. P. Lee, W. Gao, and X. Zhang, "Speckle properties of the logarithmically transformed signal in optical coherence tomography," *J. Opt. Soc. Am. A* **28**(4), 517–522 (2011).
12. M. Pircher et al., "Speckle reduction in optical coherence tomography by frequency compounding," *J. Biomed. Opt.* **8**(3), 565–570 (2003).
13. A. E. Desjardins et al., "Speckle reduction in OCT using massively-parallel detection and frequency-domain ranging," *Opt. Express* **14**(11), 4736–4745 (2006).
14. M. Szkulmowski et al., "Efficient reduction of speckle noise in optical coherence tomography," *Opt. Express* **20**(2), 1337–1359 (2012).
15. L. Duan et al., "Single-shot speckle noise reduction by interleaved optical coherence tomography," *J. Biomed. Opt.* **19**(12), 120501 (2014).
16. B. Sander et al., "Enhanced optical coherence tomography imaging by multiple scan averaging," *Br. J. Ophthalmol.* **89**(2), 207–212 (2005).
17. Y. Winetraub et al., "Upper limit for angular compounding speckle reduction," *Appl. Phys. Lett.* **114**(21), 211101 (2019).
18. J. Zhao et al., "Angular compounding for speckle reduction in optical coherence tomography using geometric image registration algorithm and digital focusing," *Sci. Rep.* **10**(1), 1893 (2020).
19. M. Pircher et al., "Measurement and imaging of water concentration in human cornea with differential absorption optical coherence tomography," *Opt. Express* **11**(18), 2190–2197 (2003).
20. 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).
21. 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).
22. D. C. Fernández, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Opt. Express* **13**(25), 10200–10216 (2005).
23. Z. Jian et al., "Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform," *Opt. Express* **18**(2), 1024–1032 (2010).
24. L. Fang et al., "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomed. Opt. Express* **3**(5), 927–942 (2012).
25. H. Chen et al., "Speckle attenuation by adaptive singular value shrinking with generalized likelihood matching in optical coherence tomography," *J. Biomed. Opt.* **23**(3), 036014, (2018).
26. P. G. Daneshmand, A. Mehridehnavi, and H. Rabbani, "Reconstruction of optical coherence tomography images using mixed low rank approximation and second order tensor based total variation method," *IEEE Trans. Med. Imaging* **40**(3), 865–878 (2021).
27. Z. Amini, H. Rabbani, and I. Selesnick, "Sparse domain gaussianization for multi-variate statistical modeling of retinal OCT images," *IEEE Trans. Image Process.* **29**, 6873–6884 (2020).
28. S. Adabi et al., "A learnable despeckling framework for optical coherence tomography images," *J. Biomed. Opt.* **23**(1), 016013 (2018).
29. M. H. Eybposh et al., "Cluster-based filtering framework for speckle reduction in OCT images," *Biomed. Opt. Express* **9**(12), 6359–6373 (2018).
30. M. Gargesha et al., "Denoising and 4D visualization of OCT images," *Opt. Express* **16**(16), 12313–12333 (2008).
31. B. Chong and Y.-K. Zhu, "Speckle reduction in optical coherence tomography images of human finger skin by wavelet modified BM3D filter," *Opt. Commun.* **291**(Suppl. C), 461–469 (2013).
32. H. Yu, J. Gao, and A. Li, "Probability-based non-local means filter for speckle noise suppression in optical coherence tomography images," *Opt. Lett.* **41**(5), 994–997 (2016).
33. J. Aum, J. 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.* **54**(13), D43–D50 (2015).

34. K. Zhang et al., “Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising,” *IEEE Trans. Image Process.* **26**(7), 3142–3155 (2017).
35. N. Cai et al., “A ResNet-based universal method for speckle reduction in optical coherence tomography images,” in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)* (2018).
36. Y. Ma et al., “Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN,” *Biomed. Opt. Express* **9**(11), 5129–5146 (2018).
37. Y. Huang et al., “Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network,” *Opt. Express* **27**(9), 12289–12307 (2019).
38. B. Qiu et al., “Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function,” *Biomed. Opt. Express* **11**(2), 817–830 (2020).
39. N. A. Kande et al., “SiameseGAN: a generative model for denoising of spectral domain optical coherence tomography images,” *IEEE Trans. Med. Imaging* **40**(1), 180–192 (2021).
40. J.-L. Starck, F. D. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis: The Multiscale Approach*, Cambridge University Press (1998).
41. 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).
42. S. Chitchian, M. A. Fiddy, and N. M. Fried, “Denoising during optical coherence tomography of the prostate nerves via wavelet shrinkage using dual-tree complex wavelet transform,” *J. Biomed. Opt.* **14**(1), 014031 (2009).
43. F. Zaki et al., “Noise adaptive wavelet thresholding for speckle noise removal in optical coherence tomography,” *Biomed. Opt. Express* **8**(5), 2720–2731 (2017).
44. A. Ozcan et al., “Speckle reduction in optical coherence tomography images using digital filtering,” *J. Opt. Soc. Am. A* **24**(7), 1901–1910 (2007).
45. M. A. Mayer et al., “Wavelet denoising of multiframe optical coherence tomography data,” *Biomed. Opt. Express* **3**(3), 572–589 (2012).
46. R. Kafieh, H. Rabbani, and I. Selesnick, “Three dimensional data-driven multi scale atomic representation of optical coherence tomography,” *IEEE Trans. Med. Imaging* **34**(5), 1042–1062 (2015).
47. C. Cuartas-Vélez et al., “Volumetric non-local-means based speckle reduction for optical coherence tomography,” *Biomed. Opt. Express* **9**(7), 3354–3372 (2018).
48. J. M. Schmitt and A. Knüttel, “Model of optical coherence tomography of heterogeneous tissue,” *J. Opt. Soc. Am. A* **14**(6), 1231–1242 (1997).
49. X. Wang et al., “A two-step iteration mechanism for speckle reduction in optical coherence tomography,” *Biomed. Signal Process. Control* **43**, 86–95 (2018).
50. J. Duan et al., “Denoising optical coherence tomography using second order total generalized variation decomposition,” *Biomed. Signal Process. Control* **24**, 120–127 (2016).
51. J. Xu et al., “Wavelet domain compounding for speckle reduction in optical coherence tomography,” *J. Biomed. Opt.* **18**(9), 096002 (2013).
52. X. Li, S. Liang, and J. Zhang, “Acceleration of OCT signal processing with lookup table method for logarithmic transformation,” *Appl. Sci.* **9**(7), 1278 (2019).
53. P. S. Hiremath, P. T. Akkasaligar, and S. Badiger, “Speckle noise reduction in medical ultrasound images,” in *Advancements and Breakthroughs in Ultrasound Imaging*, IntechOpen (2013).
54. S. Gai et al., “Speckle noise reduction in medical ultrasound image using monogenic wavelet and Laplace mixture distribution,” *Digital Signal Process.* **72**, 192–207 (2018).
55. Z. Chen et al., “Phase-stable swept source OCT angiography in human skin using an akinetic source,” *Biomed. Opt. Express* **7**(8), 3032–3048 (2016).
56. Z. Chen et al., “Non-invasive multimodal optical coherence and photoacoustic tomography for human skin imaging,” *Sci. Rep.* **7**(1), 17975 (2017).
57. A. Aneesh et al., “Multispectral *in vivo* three-dimensional optical coherence tomography of human skin,” *J. Biomed. Opt.* **15**(2), 026025 (2010).
58. A. A. Khanifar and S. Farsiu, “Spectral-domain OCT in practice,” Retina Today 49-51, 2008, [https://retinatoday.com/articles/2008-may/0508\\_10-php](https://retinatoday.com/articles/2008-may/0508_10-php) (accessed 25 August 2020).

59. "AREDS 2 Ancillary Spectral Domain Optical Coherence Tomography Study (A2ASDOCT)," 2008-2013, <https://clinicaltrials.gov/ct2/show/NCT00734487> (accessed 27 August 2020).

**Juan J. Gómez-Valverde** is an assistant professor at the Universidad Politécnica de Madrid (UPM) and a researcher in the Biomedical Research Networking Center in bioengineering, biomaterials, and nanomedicine. He graduated in telecommunication engineering and completed a master's program in biomedical engineering at the UPM. He obtained his PhD with international mention from UPM. His main research interests are related to screening for diseases, biomedical image processing, artificial intelligence, and optical image analysis.

**Zhe Chen** received his PhD from the Medical University of Vienna in 2020. His research area includes dual modality optical coherence tomography angiography (OCTA) and photoacoustics tomography (PAT) imaging techniques. He is currently working in optical and architecture design of laser beam scanners used in the augmented reality area.

**Andrés Santos** received his MEng degree and PhD from UPM where he is currently a professor and the director of the Biomedical Image Technologies lab. His main research areas are biomedical image analysis and acquisition. He has coauthored over 90 papers in indexed journals on medical imaging and has coordinated more than 50 research projects and technology transfer contracts. He has been co-director of a master's program on biomedical technology and instrumentation.

**Wolfgang Drexler** is a full professor and since 2009 head of the Center for Medical Physics and Biomedical Engineering after having been full professor of biomedical imaging at Cardiff University, UK, from 2006 to 2009. He also spent two years at the MIT, Cambridge, from 1998 to 1999. He received the Austrian START Award in 2001 and the COGAN Award in 2007. His h-index is 72 (Scopus) and his research grant income since 2000 is €17 million.

**María J. Ledesma-Carbayo** received her MEng degree in telecommunication engineering and her PhD from UPM specializing in medical image analysis. She additionally completed two master's programs in biomedical engineering and medical physics. She is currently a full professor at UPM. She has authored over 150 publications in indexed journals and conferences. Her research interests include cardiac and pulmonary imaging, image-guided therapy, microscopy image analysis, registration, motion estimation, and compensation.

Biographies of the other authors are not available.