Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

A benchmark study of automated intra-retinal cyst segmentation algorithms using optical coherence tomography B-scans

GN Girish [a,*], VA Anima [a], Abhishek R Kothari [b], PV Sudeep [c,d], Sohini Roychowdhury [e], Jeny Rajan [a]

[a] *Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal, India*
[b] *Pink City Eye and Retina Center, Jaipur, India*
[c] *Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Surathkal, India*
[d] *Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal, India*
[e] *Department of Electrical and Computer Engineering, University of Washington, Bothell, USA*

ARTICLE INFO

ABSTRACT

(Background and objectives): Retinal cysts are formed by accumulation of fluid in the retina caused by leakages from inflammation or vitreous fractures. Analysis of the retinal cystic spaces holds significance in detection and treatment of several ocular diseases like age-related macular degeneration, diabetic macular edema etc. Thus, segmentation of intra-retinal cysts and quantification of cystic spaces are vital for retinal pathology and severity detection. In the recent years, automated segmentation of intra-retinal cysts using optical coherence tomography B-scans has gained significant importance in the field of retinal image analysis. The objective of this paper is to compare different intra-retinal cyst segmentation algorithms for comparative analysis and benchmarking purposes.
(Methods): In this work, we employ a modular approach for standardizing the different segmentation algorithms. Further, we analyze the variations in automated cyst segmentation performances and method scalability across image acquisition systems by using the publicly available cyst segmentation challenge dataset (OPTIMA cyst segmentation challenge).
(Results): Several key automated methods are comparatively analyzed using quantitative and qualitative experiments. Our analysis demonstrates the significance of variations in signal-to-noise ratio (SNR), retinal layer morphology and post-processing steps on the automated cyst segmentation processes.
(Conclusion): This benchmarking study provides insights towards the scalability of automated processes across vendor-specific imaging modalities to provide guidance for retinal pathology diagnostics and treatment processes.

1. Introduction

A retinal cyst is a fluid-filled space in the retina. Medical studies show that visual acuity can be accurately predicted from the volume of retinal cystic fluids and their relative location in the retina [1]. Also, Cystoid macular edema (CME), caused by cysts in the retinal macular region, is the leading cause of central vision loss in the world today. CME develops when excess fluid accumulates within the retinal macula, which may lead to disruption of the retinal vessel barrier owing to pathologies such as age related macular degeneration, diabetic retinopathy, retinal vein occlusion and ocular inflammation. This process of fluid accumulation in retina can reduce macular retinal function [2] and can lead to irreversible blindness globally in individuals belonging to the economically productive age-group, irrespective of gender and demographics [3]. Further, visual acuity impairment due to CME can be correlated to the volume of the cystic fluid spaces and their location in retinal tissue [1]. Thus, automated quantification of retinal pathology severity is imperative towards timely retinal diagnostics and treatment.

CME can be clinically characterized using Optical Coherence Tomography (OCT) images [4]. OCT is a noninvasive imaging modality that is widely used for resolving internal structures of biological tissues, and for visualizing cross-sectional high-resolution images of the retina [5]. OCT images are extensively utilized for diagnostic and prognostic purposes for several retinal pathologies with manifestations that impact the intra-retinal micro-structure, such as cysts, exudates and retinal disorganization (see Fig. 1). However, one primary limitation of the OCT images is the manual assessment time required for analyzing the large volumes of image data

* Corresponding author.
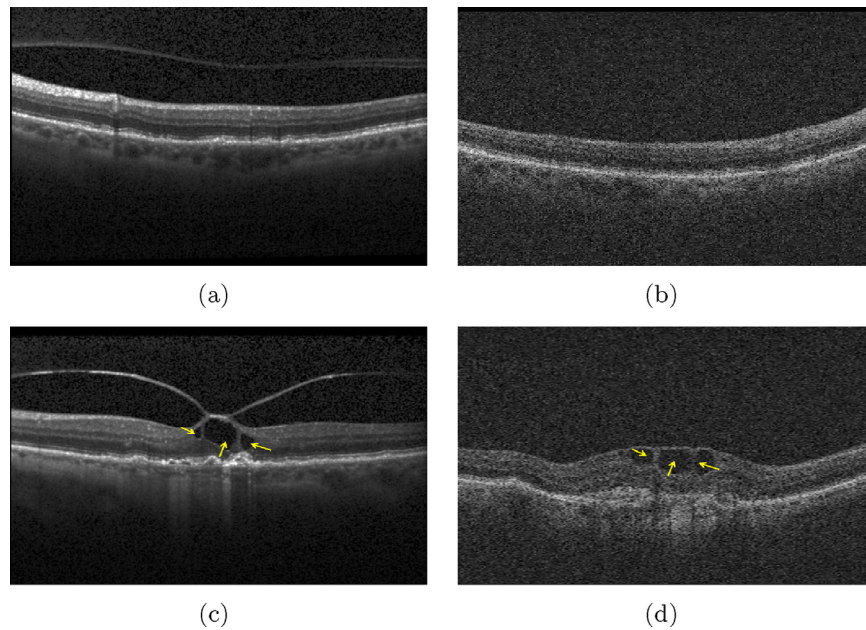*E-mail addresses:* girishanit@gmail.com, girish.cs14f07@nitk.edu.in (G. Girish).

Fig. 1. OCT image of retina from two different vendors: (a),(b) Normal retina ; (c),(d) retina with cystoid macular edema. (Yellow colored arrow shows cystic fluids). (a) and (c) obtained from Spectralis imaging system, (b) and (d) obtained from Cirrus imaging system. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
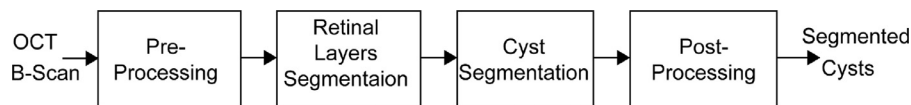

Fig. 2. Generic framework of automated intra-retinal cyst segmentation system.

per patient. This necessitates the development of automated intra-retinal and cyst segmentation/quantification methods to speed up the pathology characterization and diagnostic process. Several automation methodologies have been proposed in the recent past to address automated analysis of OCT images [1,6–10]. The key challenges posed by the OCT images to most existing methods for automated segmentation of cystic regions include pixel-level variabilities due to noise, image intensity variations, varied cyst morphology, confounding retinal structures and complex pathologies.

Based on the existing methods for automated cyst segmentation from OCT images, a generic methodological framework is proposed in Fig. 2. This framework consists of four main steps: (1) pre-processing; (2) retinal layer segmentation; (3) cyst segmentation and (4) post-processing. Since OCT images contain varying degrees of additive speckle noise, a pre-processing module is required for quality enhancement and equalization of the OCT images. Finally, the post-processing step is implemented to reduce the incorrectly segmented non-cystic regions (i.e., false positive regions). In this work, existing automated retinal cyst segmentation methods are standardized based on the work-flow shown in Fig. 2 and comparatively analyzed to evaluate the significance of the automated methods with respect to input data and output metrics.

This paper makes three key contributions. First, a modular approach to standardize existing OCT cyst segmentation methods is presented for methodological benchmarking purposes. The methodological contributions from significant automated OCT cyst segmentation methods are reviewed and comparatively discussed. Second, quantitative and qualitative analysis experiments are presented for evaluation of the existing automated OCT cyst segmentation methods. We observe that supervised OCT segmentation methods achieve higher cyst segmentation recall when compared to unsupervised approaches with degradation in segmentation precision across data sets with variable scan qualities. Third, OCT images from two different image acquisition systems are comparatively analyzed for scalability limitations owing to the image-level variabilities introduced by imaging systems. Such exhaustive analysis regarding the scalability of OCT cyst segmentation methods in terms of methodological and input data variations has not been presented so far. This work provides novel insights into the limitations of automated cyst segmentation tasks for retinal diagnostic and screening purposes.

The organization of this paper is as follows. In Section 2, the materials and evaluation metrics used for comparison study are presented. In Section 3, methods considered for comparative study are briefly reviewed. In Section 4, the experimental setup is discussed. In Section 5, the experimental results of the proposed methods are presented. Conclusions and discussion regarding the comparative assessment of the automated intra-retinal cyst segmentation along with future research directions are presented in Section 6.

2. Materials

2.1. Dataset

This work comparatively analyzes existing automated intra-retinal OCT cyst segmentation methods on the publicly available OPTIMA cyst challenge OCT dataset [11]. This dataset contains OCT scans obtained from CME subjects using four different imaging systems, namely Zeiss Cirrus, Nidek, Spectralis Heidelberg and Topcon. In this work, OCT scans from Cirrus and Spectralis image acquisition systems are analyzed, since the data sets from these systems demonstrate moderate to severe pathological features when compared to the other imaging systems. The selected OCT scans are acquired over 6×6 mm of the macula and foveal center from subjects with CME. The OCT frames represent gray-scale images

Table 1
Dataset description.

Subset of dataset	Subset of dataset	Number of volumes	Total number of B-scans
Spectralis	**Training**	**4**	**196**
	Testing	**4**	**112**
Cirrus	**Training**	**4**	**584**
	Testing	**4**	**512**

with resolution of [496 × 512] and [496 × 1026] pixels, respectively. The description of data set used in this work is presented in Table 1 and samples of OCT B-scan frames from Spectralis Heidelberg and Cirrus imaging systems are shown in Fig. 1, respectively. This dataset can be obtained from Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna. [1]

2.2. Method performance metrics

For quantitative analysis of the automated OCT cyst segmentation methods, the prediction of cystic regions from the automated methods are compared with the manually annotated Ground Truth (GT) region images, that are provided for the input data sets from two independent trained ophthalmologists. In this work, we perform segmentation of cystic regions followed by the evaluation of the number of True Positive (TP), False Positive (FP) and False Negative (FN) regions. Here, TPs are defined as actual cystic regions that are automatically detected, whereas the non-cystic regions detected as cysts by an automated algorithm are FPs, and the actual cystic regions that are undetected by an automated algorithm are FNs. The output metrics for evaluation of the automated segmentation algorithms are defined as precision (1) and recall (2).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

As an additional measure for methodological evaluation, the Dice coefficient (DC) [12] (3) analyzes the correlation of automated segmentation vs. GT per patient image stack.

$$\text{DC} = 2 \frac{|AutomatedDetection \cap GT|}{|AutomatedDetection| \cup |GT|} \tag{3}$$

3. Methods

The existing automated OCT cyst segmentation methods can be classified into two categories: semi-automated and fully automated. The semi-automated methods require manual intervention to define initial markers for each cyst. These methods are time and manual labor intensive owing to the large numbers of frames that need to be manually examined to define the markers [13,14]. Thus, fully automated segmentation methods were developed to overcome these limitations. Fig. 3 highlights this categorization of existing methods for intra-retinal cyst segmentation. The automated methods are initially categorized as 2D or 3D segmentation techniques followed by further classification into unsupervised [1,6,7,15,16] and supervised [9,10,17–22] approaches. The methodological contributions of significant existing automated OCT cyst segmentation techniques are summarized below.

Wilkins et al. [1] proposed an automated cyst segmentation method using empirically obtained global intensity threshold [1]. For reducing the speckle noise, the authors applied a fast bilateral filter on the SNR balanced and normalized OCT B-scans.

In addition, the top-bottom search mechanism with the intensity threshold was employed to identify the Retinal Pigment Epithelium (RPE)–Nerve Fiber Layer (NFL) boundaries. Next, regional thresholding was used for the rejection of FPs. In [1], the method was evaluated using 19 OCT-volumes (16 captured from eyes of patients with vitreo-retinal disease and 3 control eyes) and the mean sensitivity and specificity were reported as 0.91 and 0.96, respectively. However, cyst segmentation using empirical thresholding is not feasible in a clinical setup and thus, the post-processing steps mentioned in [6] are insufficient for generalizable removal of FPs.

Gonzalez et al. [21] used Watershed Transformation (WT) for detecting cysts followed by region merging and texture analysis. In [21], the retinal layers were initially segmented using a graph cut segmentation approach. After delimiting the region of interest (ROI) between Inner Limiting Membrane (ILM) and RPE layers, WT was applied to segment the cyst regions. After WT segmentation, the authors combined the over-flooded regions to form larger catchment basins by region merging approach and performed the post-processing steps using 14 different parameters for each region. The detected cysts were then classified with texture descriptors based on Gray Level Co-occurrence Matrix (GLCM) and Gabor filters and by using classifiers such as Naive-Bayes (NB), Support Vector Machine (SVM) or Random Forest (RF) with 10-fold cross validation. The comparative evaluation in [21] indicates that GLCM + SVM and Gabor + SVM have segmentation accuracies of 0.8293 and 0.8244, respectively. The limitations of this method include over-flooding of WT due to improper regional minima-estimation and the large number of parameters while post-processing. Hence, this method suffers from generalizability in a clinical setup.

In [17], Pilch et al. segmented the cysts with the *k*-means clustering technique and classified them using *k*-Nearest Neighbor (*k*-NN). Beforehand, the speckle noise in B-scans was removed with Bayesian estimation method, followed by implementation of an automated active contour model for retinal layer segmentation. The experiments discussed in [17], using eight OCT B-scans with 130 cysts, indicate that the retinal structures are over-segmented with *k*-means clustering method if *k* is set too low and improper delineation of segmented retinal micro-structures occurs when *k* is set too high.

Swingle et al. [18,19] proposed an automated pseudo-cyst detection algorithm, based on pixel-wise classification approach, and identified the spatial distribution of cysts in the retinal fundus images. Prior to classification, B-scans were normalized and ILM and Bruchs Membrane (BM) retinal boundaries were segmented. The authors trained an RF classifier with 14 features extracted from manually segmented data. Next, intensity profiles and image-based features were also estimated for the test data set. Finally, RF-based classification is performed on the test images, where each test-image pixel was probed for being cyst-like. Due to the presence of spurious pixels in the classification output, post-processing was performed to remove all connected components below the threshold of 30 pixels. In [18], the proposed method was evaluated on 49 OCT B-scans obtained from five subjects with Micro-cystic Macular Edema (MME) and the results demonstrated automated classification accuracy of 0.846 and DC of 0.75 obtained against two manual graders. However, this pixel-wise classification approach for cysts

---

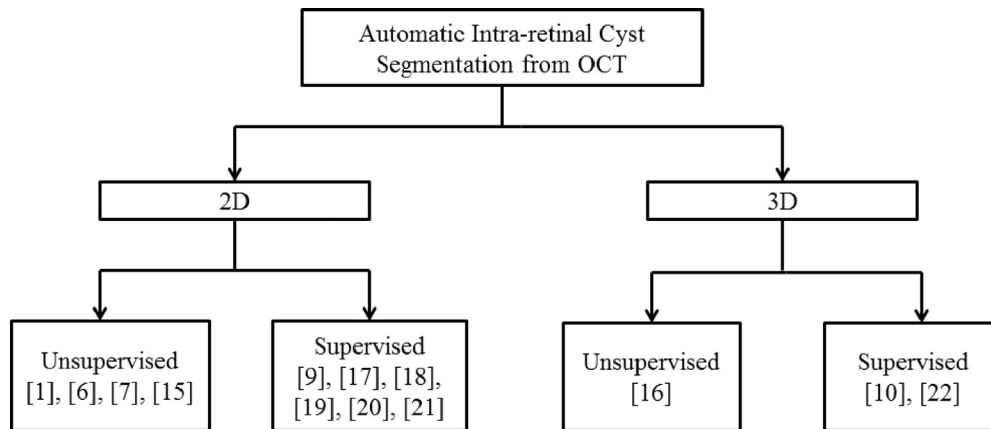[1] optima@meduniwien.ac.at.

Fig. 3. Taxonomy of segmentation approaches used for automated intra-retinal cyst segmentation.

is limited in its performance by the presence of speckle noise and shadow artifacts in OCT images.

Lang et al. [9] used both intensity and spatial features with an RF classifier for classifying pseudo-cyst pixels from OCT B-scans. Initially B-scans were normalized, followed by extraction of eighteen features, of which sixteen were intensity based and two were spatial position based features. The system was trained on the features extracted from manually segmented data. After extracting features for test data, the RF classifier was used with 60 trees, each with minimum terminal leaf size of 10. Experimental analysis was performed using leave-one-out approach. Further, this method was assessed on images obtained from 10 MME patients and 10 control candidates. For MME data, precision of 0.85, recall of 0.79 and DC of 0.98 was reported with this algorithm. In case of control data, this method correctly labeled all the pixels.

Wieclawek et al. segmentation method involved different image processing techniques including non-linear complex diffusion and mathematical morphology operations [6]. The authors segmented the cysts from OCT B-scans using *h*-minima transform and image reconstruction by erosion and geodesic operations. The segmented regions above and below the NFL and RPE layers and the regions connected to image boundaries were removed for reducing FPs. However, in this work, the performance of the cyst segmentation method were qualitatively reported.

In [7,15], Girish et al. presented an intra-retinal cyst segmentation method using marker controlled WT algorithm. The OCT B-scans were preprocessed with Bayesian Non-Local Means filter to remove the speckle noise and the six prominent retinal layers were segmented from the denoised images using OCTSEG tool [23]. In the method, the extracted markers with *k*-means clustering algorithm was used as a source for topographical watershed segmentation. Experimental analysis on the OPTIMA cyst challenge dataset [11] showed that recall and precision rates were 0.82 and 0.77, respectively. A higher DC of 0.96 was reported with the ground-truth obtained from two manual graders.

Recently, several works have been directed towards automated 2D segmentation [20,24] and 3D volumetric segmentation of retinal cysts [10,16,22]. In case of 3D volumetric analysis of OCT, densely sampled data with higher number of C-scans are required to avoid the missing of small cyst structures. The qualitative and quantitative evaluation of these methods is presented in the following sections.

4. Experimental setup

Experiments analysis for automated segmentation of cysts using the methods described above is performed on OCT images

from Spectralis and Cirrus imaging systems, and the segmentation outcomes are comparatively evaluated. For standardization purposes, the pre-processing module for the removal of additive speckle noise from the OCT B-scans precedes the retinal layer segmentation step. For each method described in Section 3, the automated segmentation algorithm module is followed by post-processing modules. The segmentation methods are implemented in MATLAB on a standalone computer with following specifications: 64-bit operating system, Core i7 3700 CPU and RAM size 8 GB. The functional modules for the automated OCT segmentation methods are described below.

4.1. Pre-processing and retinal layers segmentation

Since the segmentation methods are sensitive to noise, all the input OCT images are denoised first. Here, we employ the recently proposed adaptive and unbiased bilateral (AUB) filter [25]. In [25], a three parameter Gamma distribution function is used to fit the noisy (observed) OCT B-scans and a maximum likelihood (ML) approach is used to estimate the parameters of the Gamma distribution, based on which, the *bias B* parameter is computed. Next, *B* is subtracted from the output of an adaptive version of the conventional bilateral filter [26]. The various parameters associated with the AUB filter are as follows: Half width of the filter = 5, geometrical parameter $\hat{\sigma}_d = 1.5$, photometric spread $\hat{\sigma}_r = 2.5\hat{\sigma}$ and $0.6\hat{\sigma}$ for Cirrus and Spectralis respectively. Where $\hat{\sigma}^2$ denotes the estimated noise variance from the given input image. The estimated bias $\hat{B}$ is empirically estimated as 32 and 15 for the Spectralis and Cirrus scans, respectively.

For cyst segmentation, retinal layer segmentation is an important pre-processing step because it provides information regarding variations in morphology and thickness of retinal layers, which leads to refined the cyst segmentation results. Among different automated retinal layer segmentation methods, the recent Iowa reference algorithm (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA) [27–29] is used in this work to obtain the eleven segmented retinal layers as shown in Fig. 4.

4.2. Cyst segmentation and post-processing

The cyst segmentation methods described in Section 3 are implemented with method specific post-processing steps as follows. For [1], we implement Wilkins cyst segmentation method with empirically computed gray level threshold of 49 and 37 for Spectralis and Cirrus scans, respectively. Post-processing is achieved with the gray level intensity standard deviation of 41, by removing regions less than 10 pixels and regions outside ROI (NFL-RPE).
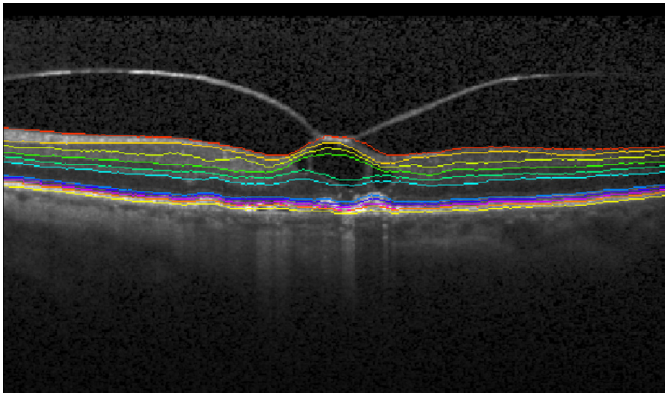
Fig. 4. OCT image of retina to visualize the order and position of the boundaries of layers. (From top: 1- internal limiting membrane (ILM); 2- retinal nerve fiber layer boundary (RNFL); 3- ganglion cell layer (GCL); 4- inner plexiform layer (IPL); 5 - outer plexiform layer (OPL); 6- outer nuclear layer (ONL); 7- Boundary of myoidand ellipsoid of inner segments (BMEIS); 8- Inner Segment/Outer Segment (IS/OS) junction; 9- Outer Photoreceptor (OPR); 10- Retinal Pigment Epithelium (RPE); 11- Bruch's Membrane Complex (BMC). (The Iowa reference algorithms (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA) [27–29]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This step reduces FPs. For Gonzalez method [21], we performed the post processing after WT segmentation as follows: elongated cysts ($t_1 = 1.5; t_2 = 2; p_{occ} = 0.6$), small cysts (regions lesser than $n_{min} = 10$ pixels), cysts next to image slides, cysts above ILM and below Outer Plexiform Layer (OPL) are removed to reduce FPs followed by classification [21].

In the implementation of k-means clustering and k-NN method [17], pre-processed B-scans are automatically segmented using k-means algorithm. Optimal value of k is empirically determined as 3. Next, the gray-level pixel intensities of B-scans are clustered as hyper reflective layer, hypo-reflective layer and cystic or low intensity regions, respectively. Finally, the clusters are post-processed to remove outliers using boundary layers of retina (ILM and RPE), pseudo-cyst ratio (lesser than 10 pixels) and width to height ratio (width is 4 times greater than height). After this post-processing step, texture features are extracted using Hu-moments and k-NN classification ($k = 6$) is performed with manually segmented training data from 4 OCT volumes.

Swingleś method in [18] is implemented to extract 14 intensity features on manually segmented data from the 4 OCT volumes as follows: 2 features are obtained from the voxel intensity before and after gray-scale morphological closing operation and 9 features are retrieved after Gaussian filtering at various scales. Next, 2 features are extracted by Laplacian and Laplacain of Gaussian of the image and a final feature is the relative distance of intra-retinal cyst to the retinal boundary (ILM-BMEIS). Finally, an RF classifier is trained from the features with 60 trees and the test data is probed. Here, we ignore all connected components lesser than 30 pixels to remove spurious pixels in the post-processing step.

The method by Lang et al. [9] is implemented with 15 features that are extracted from the OCT images, followed by the aforementioned steps for training RF classifier with parameters mentioned in [9]. Next, the test data is probed to classify cyst and non-cyst pixels. To generate final segmentation results, we employed two-stage thresholding approach with probability threshold of 0.5 and 0.85, respectively. Here, we ignore all connected components lesser than 5 pixels to remove spurious pixels.

For the method by Wieclawek [6], we fixed $h = 3$ and $h = 4$ to compute h-minima transform on Spectralis and Cirrus scans, respectively. Next, h-minima transformed B-scans are thresholded to obtain the segmented results. The segmented regions above and below NFL and RPE layers and regions connected to image boundaries are removed to reject FPs.

In [7], the B-scans are initially segmented by k-means clustering. We used $k = 3$ for k-means clustering because hypo-reflective, hyper-reflective and low-intensity cystic regions can efficiently be grouped into 3 different clusters. Next, markers are created from the center pixel of each region in the minimum centroid cluster, and these markers are used as the sources to segment the image with topography-based WT. In the post-processing step, we remove FPs by eliminating regions with the following features: region-wise intensity threshold greater than 105, elongated regions with width 4 times greater than height, pseudo-cysts of area less than 10 pixels.

5. Results and analysis

The quantitative and qualitative comparative assessment of the automated cyst segmentation methods is presented below. Since low precision is indicative of FPs (over-detection) and low recall is indicative of missing patients with abnormalities (under-detection), high precision and high recall values are desired for an ideal automated cyst segmentation method.

5.1. Quantitative assessment

Table 2 and 3 demonstrate precision and recall for the methods under analysis against GT provided by Grader 1 (G1), GT provided by Grader 2 (G2) and their intersections (G1 ∩ G2) on the Spectralis and Cirrus scans, respectively. Here, we observe that cyst segmentation obtained with unsupervised methods, such as the marker controlled watershed transform proposed by Girish et al. [7] and mathematical morphology based approach proposed by Wieclawek [6], achieve higher recall across input dataset variabilities (i.e., mean recall of 0.37-Girish, 0.38-Wieclawek and 0.38-Girish, 0.59-Wieclawek on Spectralis and Cirrus scans, respectively, with G1). However, segmentation precision varies largely with the image quality (i.e., mean precision of 0.33-Girish, 0.35-Wieclawek and 0.09-Girish, 0.09-Wieclawek on Spectralis and Cirrus scans, respectively, with G1).

Among the supervised cyst segmentation methods, Swingle et al. [18] method outperforms other supervised methods in terms of recall while preserving marginal precision rates across dataset variabilities. Also, trends in degradation of precision with respect to image quality, similar to that of the unsupervised methods, is observed. Thus, from Tables 2 and 3, we infer that supervised methods outperform unsupervised methods in terms of overall precision and recall. Also, texture-based supervised methods such as GLCM, Gabor and Hu-moments are found to be efficient discriminators for cystic regions. Additionally, from Tables 2 and 3, we observe that the variations in performances with respect to manual graders due to inter-grader variability [11].

Further, comparative assessment of cyst segmentation methods on a sample Spectralis OCT B-scan are presented in Fig. 1. The original B-scan and GT provided by G1 (see the marked region in red colour) are shown in Figs. 5(a) and (b), respectively. We observe from Fig. 5(c) that Wilkins method has failed to identify some cysts due to intensity variations, and it has also failed to delineate the cysts up to their actual boundary. In Fig. 5(d), the Wieclawek method shows over-segmentation by considering consecutive cysts as a single entity in an image. Here, the boundaries between cysts are missed and FPs in the ONL layer are observed.

In most existing works, intra-retinal cysts segmentation by automated algorithms is delimited between the ILM and RPE layers. However, the method in [7] considers the region between ILM and OPL to be the most susceptible to intra-retinal cysts, thereby leading to accurate cyst segmentation in Fig. 5(e). However, over seg-

Table 2

Mean ± Standard Deviation of Precision and Recall of compared methods on Spectralis OCT scans.

Method	G1		G2		G1 ∩ G2	
	Precision	Recall	Precision	Recall	Precision	Recall
Wilkins et al. [1][a]	0.2415 ± 0.2573	0.4584 ± 0.3073	0.2429 ± 0.2607	0.4878 ± 0.3477	0.2423 ± 0.2580	0.4620 ± 0.3259
Wieclawek [6][a]	0.3578 ± 0.2210	0.3863 ± 0.3341	0.3700 ± 0.2337	0.4282 ± 0.3533	0.3807 ± 0.2483	0.4204 ± 0.3383
Girish et al. [7][a]	0.3300 ± 0.2997	0.3758 ± 0.2625	0.3458 ± 0.3163	0.4354 ± 0.3173	0.3373 ± 0.3075	0.3990 ± 0.2788
GLCM+NB [21][b]	0.5354 ± 0.4403	0.6563 ± 0.4619	0.5373 ± 0.4417	0.6667 ± 0.4714	0.5463 ± 0.4333	0.6458 ± 0.4533
GLCM+SVM [21][b]	0.5625 ± 0.4270	0.4688 ± 0.3840	0.5875 ± 0.4250	0.5521 ± 0.4682	0.5625 ± 0.4270	0.4688 ± 0.3840
GLCM+RF [21][b]	0.5800 ± 0.4214	0.5208 ± 0.4429	0.3239 ± 0.3761	0.4167 ± 0.5000	0.2656 ± 0.3078	0.2604 ± 0.3234
Gabor+NB [21][b]	0.4000 ± 0.3031	0.3750 ± 0.2846	0.3997 ± 0.2986	0.3646 ± 0.2792	0.3889 ± 0.2940	0.3542 ± 0.2753
Gabor+SVM [21][b]	0.4167 ± 0.2887	0.3542 ± 0.2753	0.3505 ± 0.2865	0.3229 ± 0.2554	0.3824 ± 0.2893	0.3437 ± 0.2730
Gabor+RF [21][b]	0.2667 ± 0.3485	0.1979 ± 0.2342	0.2281 ± 0.2818	0.1979 ± 0.2342	0.0833 ± 0.1667	0.0104 ± 0.0208
Pilch et al. [17][b]	0.4067 ± 0.4611	0.5833 ± 0.4410	0.4125 ± 0.4686	0.6250 ± 0.4787	0.4067 ± 0.4611	0.5833 ± 0.4410
Swingle et al. [18][b]	0.3500 ± 0.2428	0.8270 ± 0.2617	0.3353 ± 0.2537	0.8140 ± 0.2852	0.3493 ± 0.2677	0.6994 ± 0.3012
Lang et al. [9][b]	0.8000 ± 0.1897	0.5388 ± 0.3403	0.7355 ± 0.2742	0.5053 ± 0.3519	0.7880 ± 0.2114	0.4816 ± 0.3134

[a] Supervised methods.
[b] Unsupervised methods

Table 3

Mean ± Standard Deviation of Precision and Recall of compared methods on Cirrus OCT scans.

Method	G1		G2		G1 ∩ G2	
	Precision	Recall	Precision	Recall	Precision	Recall
Wilkins et al. [1][a]	0.1352 ± 0.1706	0.1959 ± 0.3381	0.1344 ± 0.1704	0.2383 ± 0.4133	0.1352 ± 0.1706	0.2047 ± 0.3548
Wieclawek [6][a]	0.0968 ± 0.0889	0.5989 ± 0.4192	0.0837 ± 0.0712	0.6234 ± 0.4449	0.0944 ± 0.0848	0.6048 ± 0.4262
Girish et al. [7][a]	0.0941 ± 0.1349	0.3889 ± 0.1972	0.0899 ± 0.1282	0.4162 ± 0.1704	0.0923 ± 0.1319	0.3167 ± 0.2113
GLCM+NB [21][b]	0.3783 ± 0.2766	0.3238 ± 0.2695	0.3529 ± 0.2716	0.2779 ± 0.2192	0.3917 ± 0.2833	0.2660 ± 0.2282
GLCM+SVM [21][b]	0.5618 ± 0.4484	0.2772 ± 0.2296	0.6610 ± 0.4552	0.2209 ± 0.1928	0.6610 ± 0.4552	0.2209 ± 0.1928
GLCM+RF [21][b]	0.5101 ± 0.3821	0.1475 ± 0.1556	0.6643 ± 0.4508	0.0838 ± 0.0899	0.6875 ± 0.4732	0.0619 ± 0.0633
Gabor+NB [21][b]	0.2202 ± 0.1381	0.6150 ± 0.2656	0.2143 ± 0.1390	0.6466 ± 0.2107	0.2122 ± 0.1378	0.6693 ± 0.1920
Gabor+SVM [21][b]	0.1934 ± 0.1166	0.5628 ± 0.3139	0.1852 ± 0.1127	0.6155 ± 0.2677	0.1714 ± 0.1002	0.6064 ± 0.2369
Gabor+RF [21][b]	0.1446 ± 0.1371	0.6734 ± 0.0752	0.1236 ± 0.2071	0.6344 ± 0.0975	0.1478 ± 0.1400	0.6265 ± 0.1975
Pilch et al. [17][b]	0.0936 ± 0.0939	0.7674 ± 0.2539	0.0927 ± 0.0899	0.7735 ± 0.2327	0.0918 ± 0.0939	0.7565 ± 0.2504
Swingle et al. [18][b]	0.1784 ± 0.1821	0.9452 ± 0.0997	0.1642 ± 0.1622	0.9995 ± 0.1183	0.1758 ± 0.1966	0.9576 ± 0.0736
Lang et al. [9][b]	0.6428 ± 0.4738	0.0822 ± 0.0888	0.5000 ± 0.5773	0.0710 ± 0.0910	0.6250 ± 0.4787	0.0991 ± 0.1040

[a] Supervised methods.
[b] Unsupervised methods

mentation of consecutive cystic structures and failure to segment some true cystic regions due to intensity variation are some of the limitations to the method in [7]. Also, we observe a FP region due to the incorrect layer segmentation in Fig. 5(e). We observe from Fig. 5(f) that the Gonzalez method introduces an additional cyst (on the left side of the retina) and fails to segment consecutive cysts due to improper regional minima-estimate and over-flooding. Although the Pilch method is able to detect both micro and macro-cysts, it fails to delineate the cysts to their actual boundary as shown in Fig. 5(g)). In case of Swingle method and Lang method, all the identified consecutive cysts are merged into a large cysts (see Fig. 5 (h) and (i), respectively) due to the lack of boundary discrimination around each individual cyst.

Besides precision and recall, DC is also evaluated to validate the experimental results and plotted in Figs. 6 and 7 for Spectralis and Cirrus scans, respectively. In these box plots, the line inside the box indicates the median value and as a rule of thumb, any segmentation method performs well if the corresponding median in the box plot is high. Here, the range of the box gives the spread of the DC-values that are computed from different patient image volumes. It can be noticed from Figs. 6 and 7 that the Lang method has the highest median values for both Spectralis and Cirrus scans, respectively.

Further analysis of the impact intra-retinal layer segmentation errors on the automated cyst segmentation methods is presented using manual and automatic intra-retinal layer segmentation methods in the supplementary material (Table 1). In addition, we analyze the accuracy of different cyst segmentation algorithms on segmenting cysts located in the inner sub-retinal layers. These results are included in supplementary material (Table 2), which de-scribes the location of cysts and the corresponding performance of automated segmentation. Additionally, we analyze normal B-scan images to assess the performance of automated methods, and their results are presented in supplementary material (Table 3).

5.2. Qualitative assessment

In this sub-section, we discuss the qualitative limitations of the aforementioned methods to segment cysts from OCT images. Future works that address these limitations can significantly enhance the accuracy of automated cyst segmentation and retinal pathology diagnostics.

First, for the Wilkins method, the number of FPs are high and precision is low due to the lack of proper post-processing. For e.g., retinal blood vessel shadows and elongated cysts are not removed here. This method also suffers from manual thresholding, thus leading to low generalizability. We observe that for a high threshold value, the number of FPs tremendously increases while for a low threshold value the number of FNs increase. Similar observations are found for the Wieclawek method. The Girish method produces a high recall rate across the datasets under analysis. However, segmentation precision using this method is found to be sensitive to image noise, which in turn influences the marker generation process, that may lead to several FPs. We observe that the Cirrus scans are more noisy and they have a low mean precision value when compared to the Spectralis scans. With the Gonzalez method, the Bayesian classifier outperforms SVM and RF classifiers. This indicates that the selection of classifier plays an important role for cyst segmentation. Also, GLCM features are found to produce better accuracy than Gabor filters. From this analysis, we infer
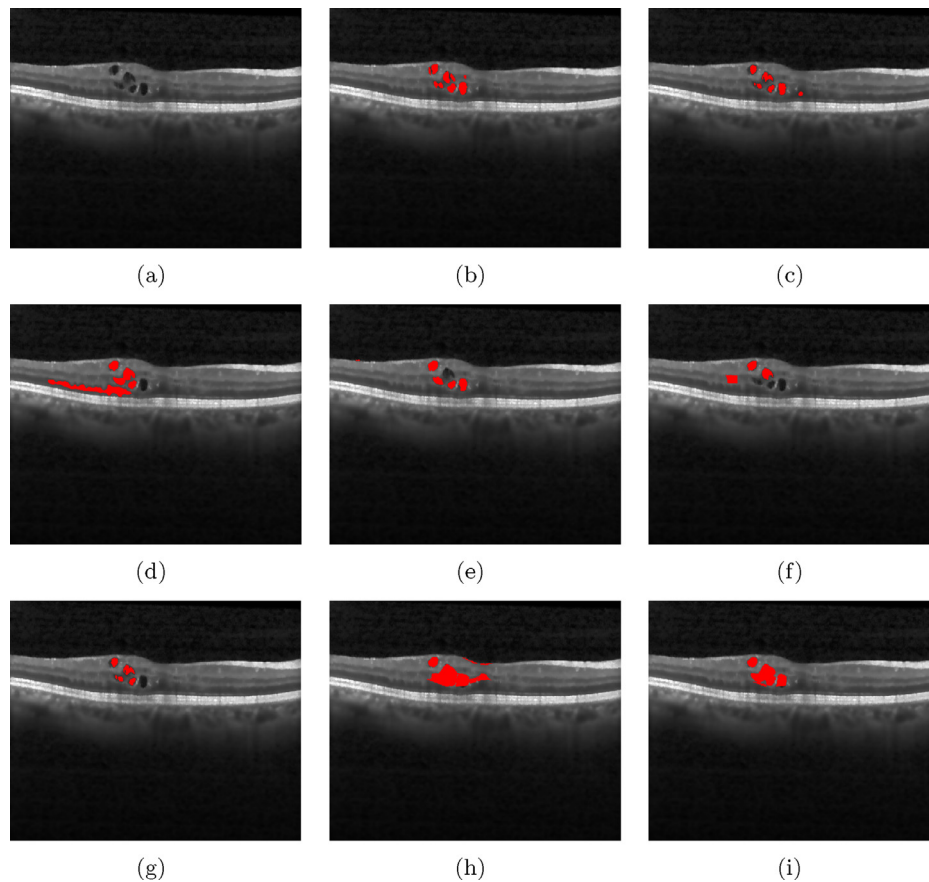
Fig. 5. Results of different automated intra-retinal cyst segmentation methods on Spectralis scans against the GT from G1 (a) Original B-scan, (b) GT (c) Wilkins Method [1], (d) Wieclawek Method [6], (e) Girish Method [7], (f) Gonzalez Method [21], (g) Pilch Method [17], (h) Swingle Method [18], (i) Lang Method [9]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
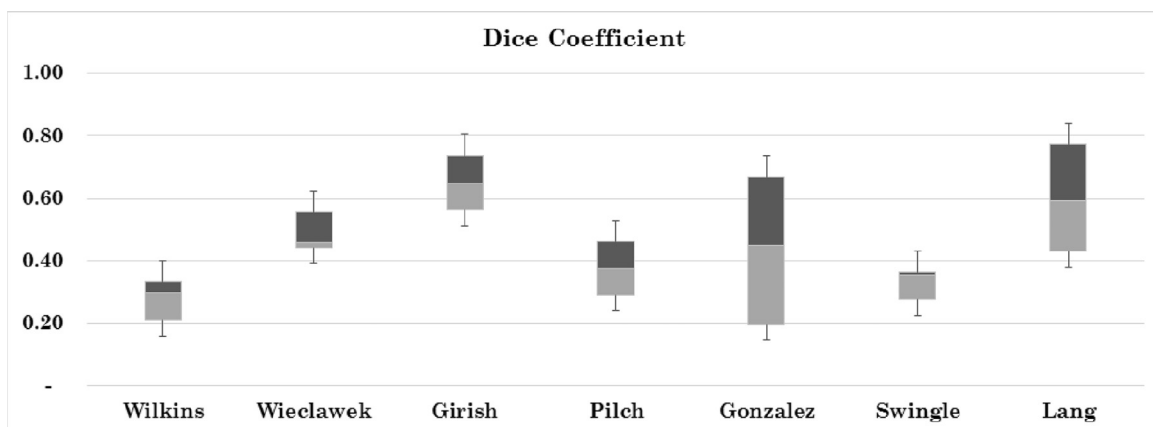

Fig. 6. Whisker Box plot for the Dice Correlation Coefficient against intersection of both the Graders on Spectralis dataset for Wilkins Method [1] (0.30), Wieclawek Method [6] (0.46), Girish Method [7] (0.65), Pilch Method [17] (0.38), Gonzalez Method [21] (0.45), Swingle Method [18] (0.35), Lang Method [9] (0.59),) .

that texture-based features have a significant role in discriminating cystic structures from other regions in OCT B-scans.

Further, we observe from the recall values with Pilch method, that this method performs well in the low intensity regions even though it is more prone to generating FPs in the presence of pixel intensity variations and blood vessel shadows. Our experiments with the Pilch method also demonstrates that Hu-moment features extracted from the segmented regions are sufficient for separating cysts from non-cysts. Besides, we observe that several factors including blood vessel shadows, epi-retinal membrane folds and distortions in layers can affect the method by Swingle et al. [18] and Lang et al. [9]. Also, we observe that the empirically se-

lected threshold values for classifiers yields better segmentation accuracies when compared to the default threshold value of 0.5. Finally, we observe that the three stage post-processing of classifier output with the method proposed in [9] leads to higher precision by reducing FPs when compared to that in the Swingle method.

Apart from cyst segmentation, in this work, we analyze the impact of the denoising module on the overall cyst segmentation process. The DC box plots charts in Figs. 8 and 9 show the variations in precision and recall for different segmentation methods before and after denoising. Here, we observe that DC values gradually increase when images are segmented after denoising. Thus, we infer
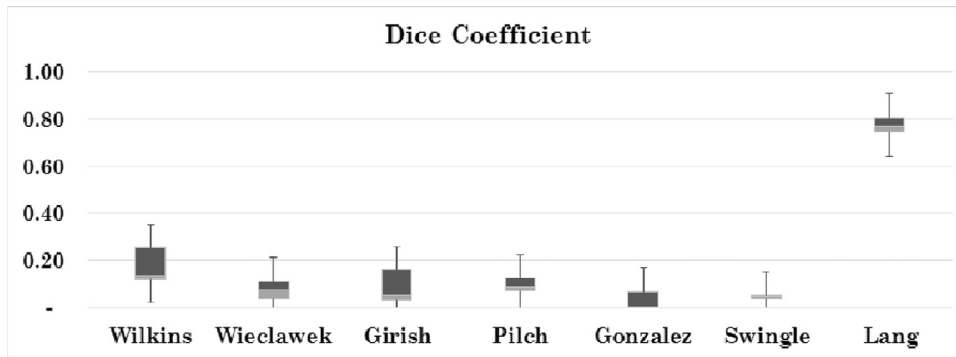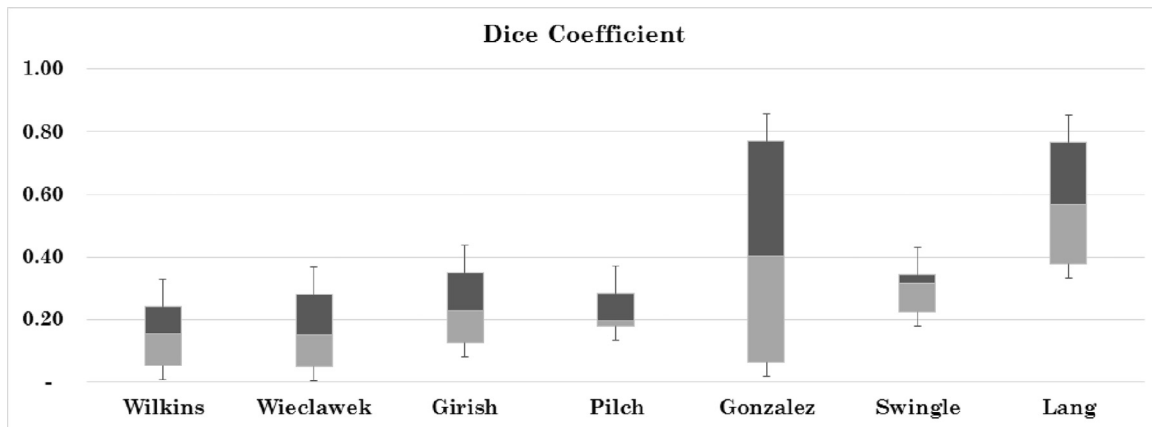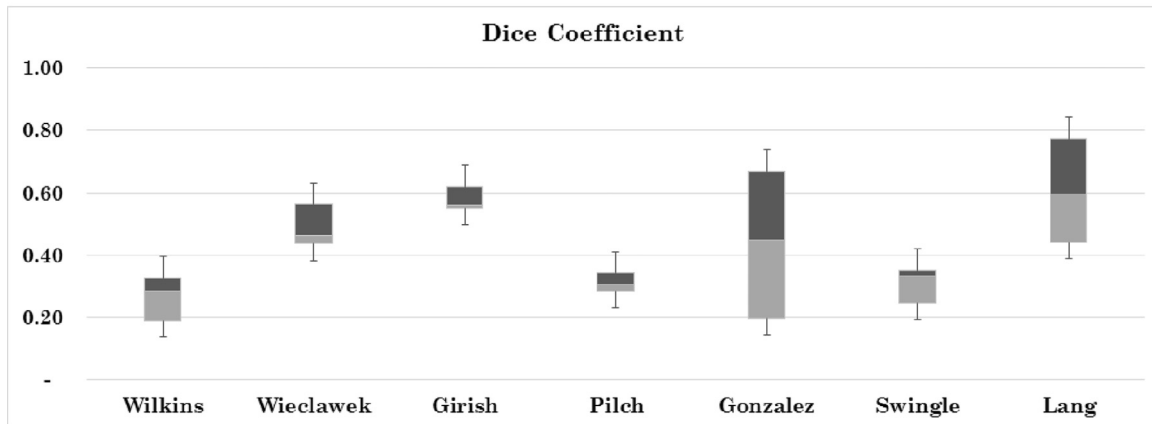
Fig. 7. Whisker Box plot for the Dice Correlation Coefficient against intersection of both the Graders on Cirrus dataset for Wilkins Method [1] (0.14), Wieclawek Method [6] (0.08), Girish Method [7] (0.05), Pilch Method [17] (0.12), Gonzalez Method [21] (0.00), Swingle Method [18] (0.05), Lang Method [9] (0.77),) .

(a)

(b)

Fig. 8. Results of DC in effect of denoising process on different automated intra-retinal cyst segmentation methods on Spectralis scans against the ground truth obtained from Grader 1 ; (a) Before denoising, (b) after denoising.

that the presence of image noise significantly distorts the segmentation of intra-retinal cystic regions.

In our experiments, we observe that some of the automated segmentation methods are effective only in detecting the cyst locations but not at delineation of the actual lesion boundary, while other methods are sensitive to method parameters, thus lacking generalizability. We know that a very important goal of any automated cyst segmentation technique should be to segment the cyst boundary accurately for volumetric assessment and follow-up treatment protocols. This work is aimed at providing guidance for

the selection of an optimal automated intra-retinal layer and cyst segmentation method, given a specific clinical/research set-up. Our analysis demonstrates that for clinical applications, the most significant metrics for selection of an automated segmentation method are: high recall rate and DC, followed by high precision and low-variability across vendor-specific imaging systems.

6. Discussion and conclusion

This paper presents a comparative assessment of existing automated intra-retinal cyst segmentation methods on OCT B-scans.
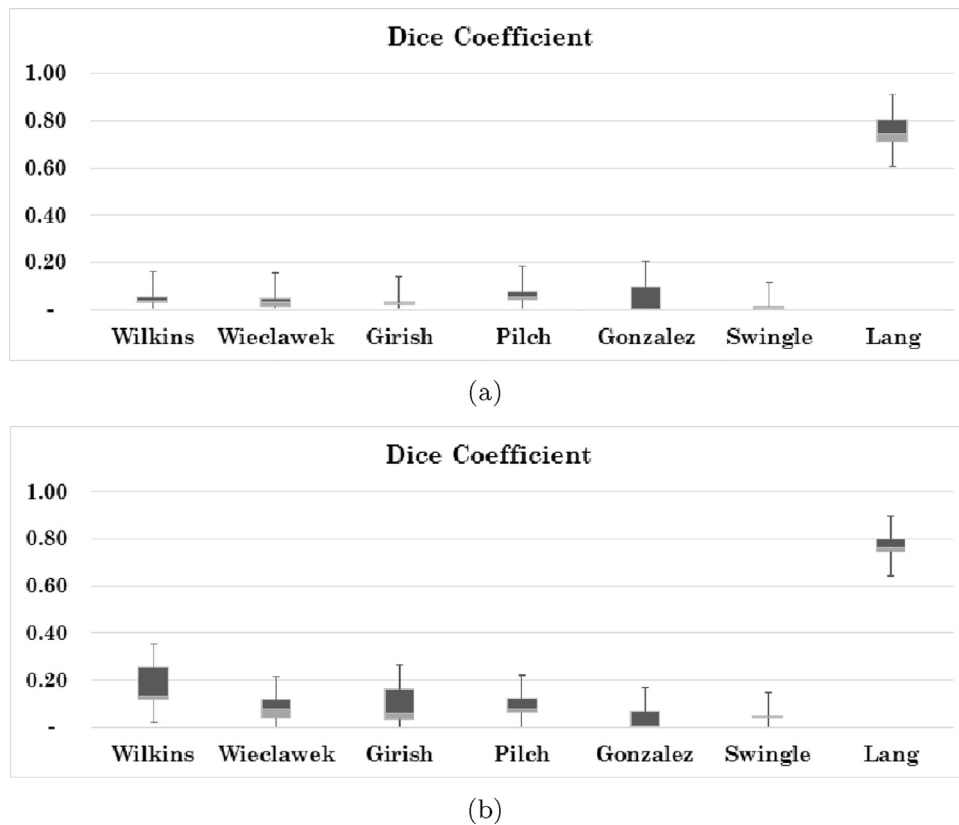
Fig. 9. Results of DC in effect of denoising process on different automated intra-retinal cyst segmentation methods on Cirrus scans against the ground truth obtained from Grader 1 ; (a) Before denoising, (b) After denoising.

Our standardized methodological modules and cyst segmentation experiments demonstrate that variability factors such as pixel intensity variations, noise, blood vessel shadows and retinal layer distortions can impact the automated cyst segmentation accuracies. In addition, pre-processing and post-processing steps are found to play a vital role in automated cyst segmentation processes.

An efficient cyst segmentation technique is one that is robust against the aforementioned variability factors and that is capable of performing cyst identification and delineation with least errors. In this work, the performances of a variety of supervised and unsupervised methods are compared using OCT images from Spectralis and Cirrus acquisition systems and the cyst segmentation results are quantitatively and qualitatively analyzed. Such exhaustive analysis of existing OCT cyst segmentation methods has not been presented so far. Our analysis shows that the performance of the unsupervised methods highly depends on the quality of OCT B-scans and the supervised methods outperform unsupervised methods, even if the quality of the OCT B-scan is low. However, the limitations posed by supervised methods include: over-segmentation in the presence of several consequent cysts, and the need for training data. Future works will be directed towards algorithm fine-tuning and the application of deep learning methods for a robust supervised cyst segmentation system that can efficiently discriminate cyst and non-cyst pixels.

We observe that several factors must be considered while designing an automated cyst segmentation technique. The desirable features to be included in prospective algorithms would include (but not limited to):

- Image denoising and retinal layer segmentation.
- Removal of blood vessel and hard exudate shadow artifacts prior to segmentation.

- Detection of candidate regions and accurate delineation of lesion boundaries.
- Post-processing operations for removal of false positive candidate regions.

Acknowledgments

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.cmpb.2017.10.010.

References

[1] G.R. Wilkins, O.M. Houghton, A.L. Oldenburg, Automated segmentation of intraretinal cystoid fluid in optical coherence tomography, Biomed. Eng. IEEE Trans. 59 (4) (2012) 1109–1114.
[2] S.R. Irvine, A newly defined vitreous syndrome following cataract surgery, Am. J. Ophthalmol. 36 (5) (1953) 600.
[3] S. Sivaprasad, B. Gupta, R. Crosby-Nwaobi, J. Evans, Prevalence of diabetic retinopathy in various ethnic groups: a worldwide perspective, Surv. Ophthalmol. 57 (4) (2012) 347–370.
[4] P.G. Garg, et al., Cystoid macular edema, 2014.
[5] D. Huang, E.A. Swanson, C.P. Lin, Schuman, et al., Optical coherence tomography, Science 254 (5035) (1991) 1178–1181.
[6] W. Wieclawek, Automatic cysts detection in optical coherence tomography images, in: Mixed Design of Integrated Circuits & Systems (MIXDES), 2015 22nd International Conference, IEEE, 2015.

[7] G.N. Girish, A.R. Kothari, J. Rajan, Automated segmentation of intra-retinal cysts from optical coherence tomography scans using marker controlled watershed transform, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, doi:10.1109/EMBC.2016. 7590943.

[8] S. Roychowdhury, D. Koozekanani, S. Radwan, K. Parhi, Automated localization of cysts in diabetic macular edema using optical coherence tomography images, in: Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE, IEEE, 2013.

[9] A. Lang, A. Carass, E.K. Swingle, O. Al-Louzi, P. Bhargava, S. Saidha, H.S. Ying, P.A. Calabresi, J.L. Prince, Automatic segmentation of microcystic macular edema in oct, Biomed. Opt. Express 6 (1) (2015) 155–169.

[10] L. Zhang, W. Zhu, F. Shi, H. Chen, X. Chen, Automated segmentation of intraretinal cystoid macular edema for retinal 3D oct images with macular hole, Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on, IEEE, 2015.

[11] Optima cyst segmentation challenge, (https://optima.meduniwien.ac.at/ research/challenges/) 2015.

[12] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.

[13] D.C. Fernandez, Delineating fluid-filled region boundaries in optical coherence tomography images of the retina, Med. Imaging IEEE Trans. 24 (8) (2005) 929–945.

[14] Y. Zheng, J. Sahni, C. Campa, A.N. Stangos, A. Raj, S.P. Harding, Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain optical coherence tomography images of the retina, Am. J. Ophthalmol. 155 (2) (2013) 277–286, doi:10.1016/j.ajo.2012.07.030.

[15] G.N. Girish, A.R. Kothari, J. Rajan, Marker controlled watershed transform for intra-retinal cysts segmentation from optical coherence tomography b-scans, Pattern Recognit. Lett., (2017), doi:10.1016/j.patrec.2017.08.031.

[16] X. Chen, M. Niemeijer, L. Zhang, K. Lee, M.D. Abràmoff, M. Sonka, Three-dimensional segmentation of fluid-associated abnormalities in retinal oct: probability constrained graph-search-graph-cut, Med. Imaging IEEE Trans. 31 (8) (2012) 1521–1531.

[17] M. Pilch, K. Stieger, Y. Wenner, et al., Automated segmentation of pathological cavities in optical coherence tomography scans, Invest. Ophthalmol. Vis. Sci. 54 (6) (2013) 4385–4393.

[18] E.K. Swingle, A. Lang, A. Carass, H.S. Ying, P.A. Calabresi, J.L. Prince, Microcystic macular edema detection in retina oct images, SPIE Medical Imaging, International Society for Optics and Photonics, 2014.

[19] E.K. Swingle, A. Lang, A. Carass, O. Al-Louzi, S. Saidha, J.L. Prince, P.A. Calabresi, Segmentation of microcystic macular edema in cirrus oct scans with an exploratory longitudinal study, SPIE Medical Imaging, International Society for Optics and Photonics, 2015.

[20] S.J. Chiu, M.J. Allingham, P.S. Mettu, S.W. Cousins, J.A. Izatt, S. Farsiu, Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema, Biomed. Opt. Express 6 (4) (2015) 1172–1194.

[21] A. Gonzalez, B. Remeseiro, M. Ortega, M. Penedo, P. Charlon, Automatic cyst detection in oct retinal images combining region flooding and texture analysis, Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on, 2013, doi:10.1109/CBMS.2013.6627825.

[22] G. Quellec, K. Lee, M. Dolejsi, M.K. Garvin, M.D. Abràmoff, M. Sonka, Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in sd-oct of the macula, Med. Imaging IEEE Trans. 29 (6) (2010) 1321–1330.

[23] M. Mayer, J. Hornegger, C.Y. Mardin, R.-P. Tornow, Retinal nerve fiber layer segmentation on FD–OCT scans of normal subjects and glaucoma patients, Biomed. Opt. Express 1 (5) (2010) 1358–1383.

[24] T. Schlegl, A.-M. Glodan, D. Podkowinski, S.M. Waldstein, B.S. Gerendas, U. Schmidt-Erfurth, G. Langs, Automatic segmentation and classification of intraretinal cystoid fluid and subretinal fluid in 3d-oct using convolutional neural networks, Invest. Ophthalmol. Vis. Sci. 56 (7) (2015) 5920–5920

[25] P.V. Sudeep, S. Issac Niwas, P. Palanisamy, J. Rajan, X. Yu, X. Wang, Y. Luo, L. Liu, Enhancement and bias removal of optical coherence tomography images: an iterative approach with adaptive bilateral filtering, Comput. Biol. Med. 71 (2016) 97–107.

[26] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Computer Vision, 1998. 6th International Conference on, 1998.

[27] K. Li, X. Wu, D.Z. Chen, M. Sonka, Optimal surface segmentation in volumetric images-a graph-theoretic approach, Pattern Anal. Mach. Intell. IEEE Trans. 28 (1) (2006) 119–134, doi:10.1109/TPAMI.2006.19.

[28] M.D. Abramoff, M.K. Garvin, M. Sonka, Retinal imaging and image analysis, Biomed. Eng. IEEE Rev. 3 (2010) 169–208, doi:10.1109/RBME.2010.2084567.

[29] M.K. Garvin, M.D. Abramoff, X. Wu, S.R. Russell, T.L. Burns, M. Sonka, Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images, Med. Imaging IEEE Trans. 28 (9) (2009) 1436–1447, doi:10.1109/TMI.2009.2016958.