35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

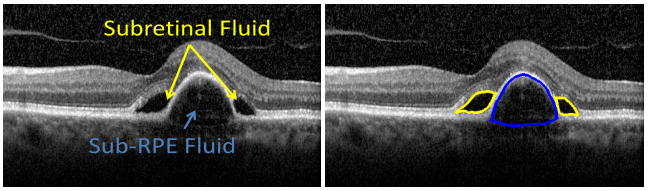
Automatic Detection of Subretinal Fluid and Sub-Retinal Pigment Epithelium Fluid in Optical Coherence Tomography Images

Weiguang Ding, Mei Young, Serge Bourgault, Sieun Lee, David A. Albiani, Andrew W. Kirker, Farzin Forooghian, Marinko V. Sarunic, Andrew B. Merkur, and Mirza Faisal Beg

Abstract— **Age-related macular degeneration (AMD) is the leading cause of blindness in developed countries. Subretinal fluid (SRF) and sub-retinal pigment epithelium (sub-RPE) fluid are signs of AMD and can be detected in optical coherence tomography images. However, manual detection and segmentation of SRFs and sub-RPE fluids are laborious and time consuming. In this paper, a novel pipeline is proposed for automatic detection of SRFs and sub-RPE fluids. First, top and bottom layers of retina are segmented using a graph cut method. Then, a Split Bregman-based segmentation method is used to segment dark regions between layers. These segmented regions are considered as potential fluid candidates, on which a set of features are generated. After that, a random forest classifier is trained to distinguish between the true fluid regions from the falsely detected fluid regions. This method shows reasonable performance in a leave-one-out evaluation using a dataset from 21 patients.**

I. INTRODUCTION

Age-related macular degeneration (AMD) is the leading cause of blindness in developed countries [1]. The "wet" or neovascular form of AMD is responsible for approximately 80-90% of severe vision loss [2]. It is characterized by an outgrowth of abnormal choroidal vessels that invade the subretinal and intraretinal spaces causing exudation and hemorrhage and leading to vision loss. Treatment with intravitreal injections of drugs that inhibit vascular endothelial growth factor (VEGF) has become the gold-standard [3]. Optical coherence tomography (OCT) has become a mainstay in ophthalmology and is extensively used in retinal clinical practice. It is vital for treatment decisions in a retinal practice for diseases such as AMD, diabetic macular edema and retinal vein occlusion. For AMD, OCT enables physicians to detect and characterize appearances of neovascular AMD within the retina, such as intraretinal fluid (IRF), subretinal fluid (SRF), and sub-retinal pigment epithelium (sub-RPE) fluid and deposit. In major clinical trials, OCT images was assessed manually to guide treatment with anti-VEGF intravitreal injections [4]. However, manual assessment has two major problems: it is time consuming and has significant interobserver discrepancy. The Comparison of Age-Related Macular Degeneration Treatment Trials (CATT) [5] showed that there is a 25% interpretation discrepancy of retinal fluid between OCT reading centers and the clinicians. OCT reading centers detected the more fluids, which means clinicians are possibly undertreating patients. The development of an automatic fluid detector can help the physicians in fast and accurate detection of IRF, SRF and sub-RPE fluid in OCT images and ultimately improve productivity and the quality of health cares.



(a) Original OCT B-scan (b) Expert manual segmentation

Fig. 1. Original OCT B-scan and expert manual segmentation: (a) subretinal fluid and sub-RPE fluid are indicated in the original B-scan image, (b) yellow represents subretinal fluid, blue represents sub-RPE fluid.

Previously, a few efforts have been made on computerized assistance and automatic detection of IRF, SRF and sub-RPE fluid and related medical image analysis problem. In [6] Split Bregman-based segmentation was used to provide raw segmentation which can be used later in expert manual segmentation. Attempt of automatic detection of IRF has been made by [7] using intensity and size-based criteria. Although its reported sensitivity and specificity are high, this thresholding based method may fail when other structures, e.g. dark layers, show similar intensity profile as fluids in more complicated situation. It also requires empirical parameter tuning given different datasets. Automatic closed-contour segmentation in ophthalmic images was proposed in [8], and used to segment fluids given their positions. In addition to fluid detection related works, many publications have been done on speckle noise reduction [9]–[12] and automatic retinal layer segmentation [9], [10], [12]–[14] for

We acknowledge funding for this project from the National Science and Engineering Research Council (NSERC) and Michael Smith Foundation for Health Research (MSFHR).

Weiguang Ding, Sieun Lee, Marinko V. Sarunic and Mirza Faisal Beg are with School of Engineering Science, Faculty of Applied Science, Simon Fraser University, 8888 University Dr, Burnaby, BC, Canada wding@sfu.ca, leeau@sfu.ca, msarunic@sfu.ca, mfbeg@sfu.ca

Mei Young, Serge Bourgault, David A. Albiani, Andrew W. Kirker, Farzin Forooghian and Andrew Merkur are with Department of Ophthalmology and Visual Sciences, University of British Columbia, 2550 Willow St, Vancouver, BC, Canada c.mei.young@gmail.com, serge.bourgault.1@ulaval.ca, awkirker@gmail.com, david_albiani@hotmail.com, farzin.forooghian@gmail.com, amerkur@gmail.com

Serge Bourgault is also with Département d'Ophtalmologie et ORL - Chirurgie cervico-faciale, Université Laval, Québec, Canada. He is supported in part by the "Programme des bourses McLaughlin du doyen de la Faculté de médecine de l'Université Laval" and the William H. Ross Fellowship in Vitreo-Retinal Excellence.

OCT images.

In this paper, we describe a novel pipeline that works well for the automatic detection of SRF and sub-RPE fluid. First potential fluids will be segmented by a combination of graph cut and Split Bregman segmentation methods. Then, each segmented potential fluid is treated as a single data point with features describing its relevant properties and labels. Based on these features and labels, a random forest classifier is then trained on the output of the prior segmentation step and is used to detect SRFs and sub-RPE fluids in new unlabeled OCT B-scans. However, we found that parameters and features that work well for detection and classification of SRF and sub-RPE fluid regions didn't work well for detecting IRF due to higher variability in its appearance. Hence, a modified algorithm dedicated to the detection of IRF is currently under development, and will be described elsewhere.

II. METHODS

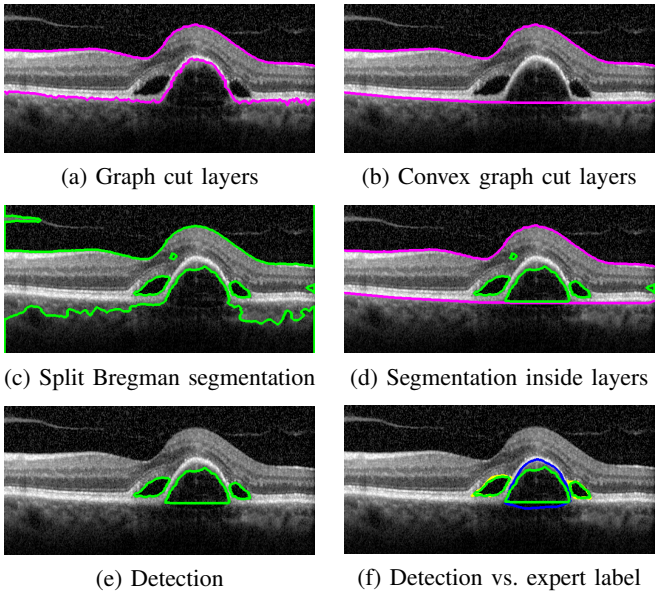
A. OCT Data Acquisition

OCT images were acquired using the Spectralis® device (Heidelberg Engineering, Heidelberg, Germany). For each OCT volume data, 25 B-scans were acquired. Each frame is an averaged 9 individual B-scans at the same location using the automatic real time (ART) eye tracking technology by Heidelberg Engineering. Each image contains 512 × 496 pixels (Figure 1) with a corresponding resolution of $4\mu\text{m}(\text{axial})\times 11\mu\text{m}(\text{lateral})$. These 25 frames are $234\mu\text{m}$ apart and are not spatially "connected", therefore they are segmented and processed separately in following analysis.

B. Segmentation and SNR Balancing

To define the region of interest of fluid detection analysis, top and bottom retina layers need to be segmented. First, original B-scans were smoothed by a bounded variation-based method [15]. The top layer, internal limiting membrane (ILM), and bottom layer, Bruch's membrane (BM), were then segmented using a graph cut-based method [13], [16] on the BV-smoothed images (Figure 2 (a)). The segmented bottom layer is usually above sub-RPE fluid, which may cause misdetection. Accordingly, convex hull [17] of the original bottom layer was taken and smoothed, and considered as a substitute for the true bottom layer (Figure 2 (b)).

A fast variational segmentation algorithm [18] was used for first round segmentation of potential fluids. This Split Bregman-based method is globally convex, fast and robust to noise. We applied the segmentation method directly on noisy OCT B-scans. Low intensity regions including background and potential fluids are separated with high intensity tissues (Figure 2 (c)). As we are only interested in detecting fluid regions between top and bottom layers, other dark regions were removed from segmented results according to already segmented layers(Figure 2 (d)). After initial segmentation, we remove any fluid regions smaller than 30 pixels. This size prior is due to that regions smaller than 30 pixels are likely artifacts caused by speckle noise.



(a) Graph cut layers (b) Convex graph cut layers

(c) Split Bregman segmentation (d) Segmentation inside layers

(e) Detection (f) Detection vs. expert label

Fig. 2. Processing pipeline: (a) graph cut segmented layers, (b) smoothed convex hull of graph cut layers, (c) Split Bregman-based segmentation, (d) segmented regions between layers, (e) automatic detection result after random forest classification, (f) automatic detection overlaid with manual segmentation ground truth.

Because the data acquired from different patients has different intensity distribution profile, normalization was performed by a signal-to-noise ratio (SNR) balancing [7] to each B-scan. We used the mean intensity of background and segmented potential fluids to represent noise level N, the mean intensity of segmented bright tissue to represents signal level S. The SNR balancing was done by $I_1 = (I_0 - N)/(S - N)$, where I_1 is the new SNR balanced image and I_0 represents the original image.

C. Sample and Feature Generation

For each fluid candidate region, a bounding rectangle was defined that it is large enough to represent the surrounding information, but not too large to include too much unrelated area. In this application, the algorithm will first try to enlarge the tight bounding box by 1.5 times but at maximum 10 pixels in each dimension. If the area of the rectangle has not reached twice the fluid size, it will enlarged to be so. Based on the bounding box and the original segmentation, shape and intensity features are extracted from each sample. Shape features include major and minor axis length, ratio of major and minor axis length, perimeter, area, ratio of perimeter and area, eccentricity, orientation, variance of the cyst height at each column. Intensity features include average intensity inside the cyst, average intensity outside the cyst, intensity difference of inside and outside, intensity variance inside, intensity kurtosis inside and intensity skewness inside. These properties constitute a 15 dimensional feature vector for each sample. For determining its label (whether a segmentation corresponds to a true fluid), we use the following definition: If given a detected region A, there exists a expert labeled true fluid region B, such that $r = \frac{area(A \cap B)}{min(area(A), area(B))}$ is larger

than a certain threshold, then A is considered to be a correct detection. Under this definition, if A or B are different in size, but one covers the other, A is still considered to be true detection.

D. Classification

Based on generated features and labels, a random forest [19] classifier was trained to classify the regions segmented as potential fluid regions. Among supervised learning algorithms, random forest is known to be fast at testing time and effective in wide range of applications [20]. At training time, n decision trees will be trained separately based on randomly selected features at each time. At testing time, to classify a single sample, majority vote of all these decision trees will be used.

We found that majority vote, which correspond to a 50% voting threshold, tended to be accurate but less sensitive on true fluid prediction. To find more true fluids at the same time allow more false positives, a modified voting procedure was applied, enabling the use of a floating threshold. Here, this floating threshold was acquired using cross validation [21] on the training set. Given a training set, it will be divided into n parts. At each iteration, one of them will be used as validation set. The other $n-1$ parts will be used to train a random forest. After these random forests were trained, they were used to do predictions on the corresponding validation sets. The threshold will be tuned to maximize the F_β measure:

$$F_\beta = (1+\beta^2) \cdot \frac{recall \cdot precision}{\beta^2 \cdot precision + recall}$$

where recall and precision are given by $recall = \frac{1}{1+FN/TP}$ and $precision = \frac{1}{1+FP/TP}$. Here, TP represents number of true positives, namely the number of detected true fluid regions. FN represents number of false negatives, which is the number of undetected true fluid regions. FP represents number of false positives, corresponding to the number of wrongly detected normal regions. In this case, larger β value results in a higher recall indicating the detector is sensitive to fluid regions; smaller β value results in a higher precision which means the detected regions are highly possible to be true fluid regions. After the desired decision threshold has been acquired by the cross validation, all the training data will be used to train a random forest to do predictions on test data. While training the random forest, we adopted the cost sensitive scheme by assigning different weight to each sample based on its label, $w = \frac{number\ of\ true\ detections}{number\ of\ false\ detections}$ is assigned on false detections and 1 on true fluids.

In the following analysis, we refer "segmentation" as the first round combined graph cut and Split Bregman-based segmentation, "classification" as the random forest classification on regions already segmented in the first round, and "detection" as the whole pipeline of "segmentation" + "classification". Under "classification", "fluid classification" means classifying segmented regions and "B-scan classification" means classifying B-scans with and without expert labeled fluids. Under "detection", "fluid detection" means the whole pipeline of segmentation and classification to find fluids and "B-scan detection" means detecting B-scans with expert labeled fluids. Here we won't distinguish "B-scan classification" and "B-scan detection", due to they are equivalent operations on all B-scans. However, "fluid classification" and "fluid detection" are different because "fluid classification" is done on all segmented regions, while "fluid detection" refers the pipeline of segmentation and classification.

III. RESULTS

The proposed novel algorithm was applied on a dataset of 21 OCT volumes from 21 subjects diagnosed with AMD. Manual OCT image segmentation was performed by three masked expert raters (AM, SB, MY), which was considered as ground truth for evaluating the automatic detection algorithm. The raters viewed the volumes on a personal computer screen and identified the SRFs and sub-RPE fluids on all frames using the segmentation editor plugin (Theodor-Boveri-Institut für Biowissenschaften, Würzburg, Germany) built for ImageJ (National Institutes of Health, Maryland). These data were automatically segmented, transformed to features, and classified according to the above described methods. For evaluation, we calculated precisions and recalls of segmentation, fluid classification, fluid detection and B-scan detection.

After the combined graph cut and Split Bregman segmentation on all of 21 data, the recall is 0.752, the precision is 0.148. Which means that 75.2% expert labeled fluids were automatically segmented in the first round segmentation and 14.8% algorithm captured regions correspond to expert segmented true fluids. In terms of B-scans, the recall is 1 and precision is 0.287.

For evaluating fluid classification, fluid detection and B-scan detection, we perform a leave-one-out test. At each iteration, we use 20 out of 21 data as training set to train a classifier and use it to do prediction on the one unused test data. Due to randomization, random forest gives slightly different results in each run. Therefore, experimentation with same parameter was run for 10 times, and their averages and standard deviations are presented.

The β value is an arbitrary choice based on the user's trade-off between recall and precision. In Figure 3, the precision-recall, β-recall and β-precision curve are plotted with errorbars for fluid classification, fluid detection and B-scan detection. According to errorbars in Figure 3, the algorithm is more random when β is small and performs more stable when β is large. In practice, however, β will rarely be assigned a small value, since β will be set to a relatively higher value to increase the recall and enable more sensitive detection. In addition, we can see, from Figure 3 (e) and (f), that when β is set to a relatively higher value, the algorithm is able to detect almost all of the B-scans containing fluids with not too many false positive B-scan predictions.

IV. DISCUSSION AND CONCLUSION

In this paper we present a novel segmentation-learning pipeline of automatic SRF and sub-RPE fluid detection.

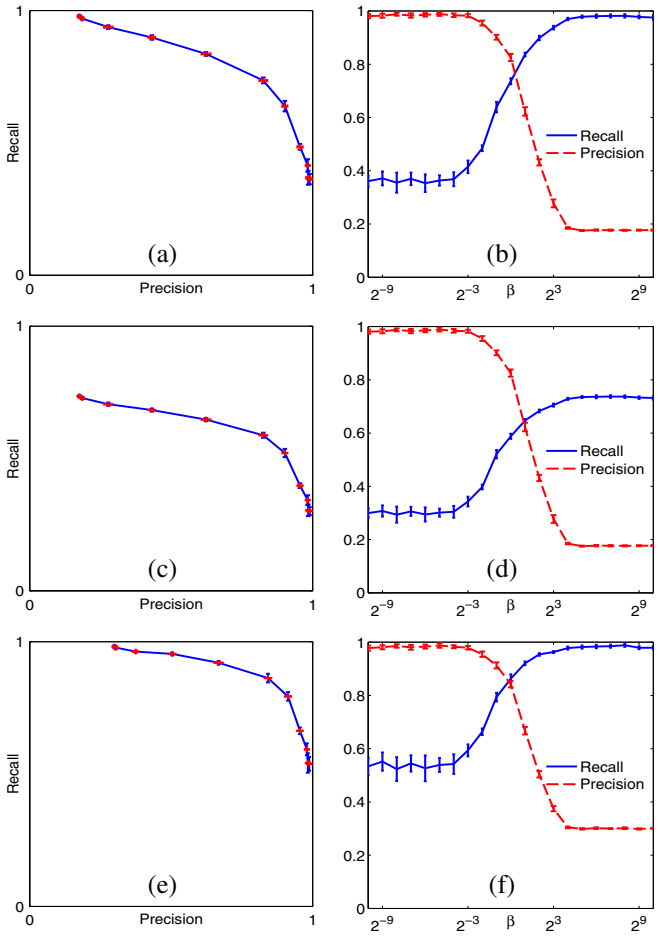


Fig. 3. Quantitative performance analysis: (a) precision-recall for fluid classification, (b) β-precision and β-recall for fluid classification, (c) precision-recall for fluid detection, (d) β-precision and β-recall for fluid detection,(e) precision-recall for B-scan detection, (f) β-precision and β-recall for B-scan detection.

Potential fluids were segmented from original B-scans and transformed to a set of descriptive features, on which random forest classification was performed to predict true fluids. Quantitative analysis was done on the performance of detection; the precision-recall curve suggests that highly sensitive and reasonably specific detection can be achieved by setting proper β value in the F_β measure. The free parameter β will usually be set to high value to ensure high recall in diagnosis.

In comparison with the previous automatic detection method [7], our method only relies on labeled training data and requires no empirical parameter tuning. It emphasizes on training the computer rather than training the human expert.

To improve the detection performance in the future, a better segmentation algorithm that is able to capture more fluids is needed. While in terms of classification, more data will increase its accuracy. The present results show that given a small dataset of 21 volumes, each containing 25 images, the algorithm performance is reasonably good. As more data is accumulated in this study, we expect both higher recall and precision for classification.

REFERENCES

[1] TY Wong, T. Wong, U. Chakravarthy, R. Klein, P. Mitchell, G. Zlateva, R. Buggage, K. Fahrbach, C. Probst, I. Sledge, et al., "The natural history and prognosis of neovascular age-related macular degeneration: a systematic review of the literature and meta-analysis.," *Ophthalmology*, vol. 115, no. 1, pp. 116, 2008.

[2] N.M. Bressler, "Age-related macular degeneration is the leading cause of blindness...," *JAMA: the journal of the American Medical Association*, vol. 291, no. 15, pp. 1900–1901, 2004.

[3] DF Martin, MG Maguire, GS Ying, JE Grunwald, SL Fine, GJ Jaffe, et al., "Ranibizumab and bevacizumab for neovascular age-related macular degeneration.," *The New England journal of medicine*, vol. 364, no. 20, pp. 1897, 2011.

[4] A.E. Fung, G.A. Lalwani, P.J. Rosenfeld, S.R. Dubovy, S. Michels, W.J. Feuer, C.A. Puliafito, J.L. Davis, HW Flynn Jr, M. Esquiabro, et al., "An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (lucentis) for neovascular age-related macular degeneration.," *American journal of ophthalmology*, vol. 143, no. 4, pp. 566, 2007.

[5] D. Dhoot and A. Yuan, "Discrepancies between clinician and reading center interpretations of oct," *Retinal Physician*, vol. 9, pp. 4–51, 2012.

[6] Y. Zheng, J. Sahni, C. Campa, A.N. Stangos, A. Raj, and S.P. Harding, "Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain optical coherence tomography images of the retina," *American Journal of Ophthalmology*, 2012.

[7] Gary R Wilkins, Odette M Houghton, and Amy L Oldenburg, "Automated segmentation of intraretinal cystoid fluid in optical coherence tomography.," *IEEE transactions on bio-medical engineering*, vol. 59, no. 4, pp. 1109–14, Apr. 2012.

[8] S.J. Chiu, C.A. Toth, C. Bowes Rickman, J.A. Izatt, and S. Farsiu, "Automatic segmentation of closed-contour features in ophthalmic images using graph theory and dynamic programming," *Biomedical Optics Express*, vol. 3, no. 5, pp. 1127–1140, 2012.

[9] D. Cabrera Fernández, H.M. Salinas, and C.A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Optics Express*, vol. 13, no. 25, pp. 10200–10216, 2005.

[10] M.K. Garvin, M.D. Abràmoff, R. Kardon, S.R. Russell, X. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-d graph search," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 10, pp. 1495–1505, 2008.

[11] P. Puvanathasan and K. Bizheva, "Interval type-ii fuzzy anisotropic diffusion algorithm for speckle noise reduction in optical coherence tomography images," *Optics express*, vol. 17, no. 2, pp. 733–746, 2009.

[12] S. Lu, C.Y. Cheung, J. Liu, J.H. Lim, C.K. Leung, and T.Y. Wong, "Automated layer segmentation of optical coherence tomography images," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 10, pp. 2605–2608, 2010.

[13] M.K. Garvin, M.D. Abramoff, X. Wu, S.R. Russell, T.L. Burns, and M. Sonka, "Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1436–1447, 2009.

[14] A. Mishra, A. Wong, K. Bizheva, and D.A. Clausi, "Intra-retinal layer segmentation in optical coherence tomography images," *Optics express*, vol. 17, no. 26, pp. 23719–23728, 2009.

[15] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*, vol. 147, Springer, 2006.

[16] K. Li, X. Wu, D.Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 1, pp. 119–134, 2006.

[17] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital image processing using MATLAB*, Pearson Education India, 2004.

[18] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split bregman method: Segmentation and surface reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1, pp. 272–293, 2010.

[19] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[20] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 161–168.

[21] C.M. Bishop et al., *Pattern recognition and machine learning*, vol. 4, springer New York, 2006.