IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 18, 2018, accepted January 3, 2019, date of publication January 21, 2019, date of current version February 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2893954

Automated 3-D Retinal Layer Segmentation From SD-OCT Images With Neurosensory Retinal Detachment

LOZA BEKALO[1], SIJIE NIU[2], XIAOJUN HE[1], PING LI[1], IDOWU PAUL OKUWOBI[1], CHENCHEN YU[1], WEN FAN[3], SONGTAO YUAN[3,4], AND QIANG CHEN[1,5], (Member, IEEE)

[1]School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[2]School of Information Science and Engineering, University of Jinan, Jinan 250022, China
[3]The First Affiliated Hospital, Nanjing Medical University, Nanjing 210029, China
[4]Jiangsu Shengze Hospital, Suzhou 214021, China
[5]Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou 350121, China

Corresponding authors: Songtao Yuan (yuansongtao@vip.sina.com) and Qiang Chen (chen2qiang@njust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671242, in part by the Key Research and Development Program of the Jiangsu Science and Technology Department under Grant BE2018131, in part by the Suzhou Industrial Innovation Project under Grant SS201759, and in part by the Open Fund Project of Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, under Grant MJUKF201706.

ABSTRACT Neurosensory retinal detachment (NRD) is a separation of the neurosensory retina from the retinal pigment epithelium (RPE) because of the subretinal fluid that can result in significant vision loss. The detachment of the neurosensory retina is known to alter the topology as well as the intensity continuity of the retinal layers. This nature of NRD makes the layer segmentation of NRD affected eyes difficult. In this paper, we presented a fully automated three-dimensional (3D) method to segment the retinal layers and NRD associated subretinal fluid from a spectral domain optical coherence tomography (SD-OCT) image. The proposed method has three phases, including a prior information model; an NRD associated subretinal fluid segmentation; and layer segmentation. The graph search and graph cut techniques were employed to segment the retinal layers and NRD associated sub-retinal fluid, respectively. To reduce the computational cost of graph-based optimization, the 'divide and merge' approach was introduced. The experiment shows that while maintaining the segmentation accuracy, the 'divide and merge' approach considerably decreases the computational cost. Our method was evaluated on 20 SD-OCT cubes diagnosed with NRD, and the results were compared with the manual segmentation results from experts. The layer evaluation showed an overall absolute surface position difference of $6.34 \pm 2.6\mu$m, which is comparable with the inter-expert variability of $6.39 \pm 5.9\ \mu$m. The segmentation result of the NRD associated sub-retinal fluid was assessed in terms of the dice coefficient and achieved means of 90.78% and 92.04% in comparison to two experts.

INDEX TERMS Graph cut, graph optimization, graph search, neurosensory retinal detachment (NRD), retinal layer segmentation.

I. INTRODUCTION

Central serous chorioretinopathy (CSC) is a vision-threatening disease among middle-aged male individuals [1]. CSC is mainly characterized by a decompensation of the retinal pigment epithelium (RPE), resulting in neurosensory retinal detachment (NRD) [2]. Segmentation of retinal layers and the boundary of NRD associated subretinal fluid are crucial for the medical assessment as well as research analysis of the retina.

Spectral domain optical coherence tomography (SD-OCT) is a non-invasive imaging modality that provides volumetric images of retinal structures with high spatial resolution. The macular centered SD-OCT image provides a clear structure of the NRD as a well-demarcated fluid region that separates the neurosensory retina from the retinal pigment epithelium (Fig. 4(A)). Despite its advantage, SD-OCT produces a relatively low signal-to-noise ratio that requires careful selection of segmentation techniques.

A number of approaches have been presented for the segmentation of retinal layers from OCT images [3]–[11]. Dufour *et al.* [5] presented an automated graph search based multi-surface segmentation method to delineate five retinal

layers. They used soft and hard smoothness constraints to add prior information from a learned model. The smoothness constraints improve the accuracy of the segmentation and increase the robustness to noise by regularizing the shape of the surface and the distance between surfaces. From retinal time-domain optical coherence tomography images, Garvin *et al.* [6] identified six surfaces using a 3D graph search. Followed by the innovation of SD-OCT imaging, this work was extended to segment seven surfaces from SD-OCT images [8]. To increase the flexibility of the smoothness constraints in handling the shape change of the layers, Garvin *et al.* [8] applied a smoothness constraint that varies across the surface. For the retinal layer segmentation from normal eyes, He *et al.* [12] developed a cascaded deep network which applies U-net to generate a probability map and regression net to learn the shape and topology of the layers'. Fang *et al.* [26] delineated nine retinal boundaries using a graph search from the probability prediction generated using FCN. The Iowa Reference Algorithms [3] is an automated algorithm developed to segment 11 retinal surfaces. The implementation of this algorithm is publicly available and the results have comparable accuracy with manual segmentation. Note that by definition, 11 retinal surfaces demarcate ten retinal layers (Fig. 1(B)).

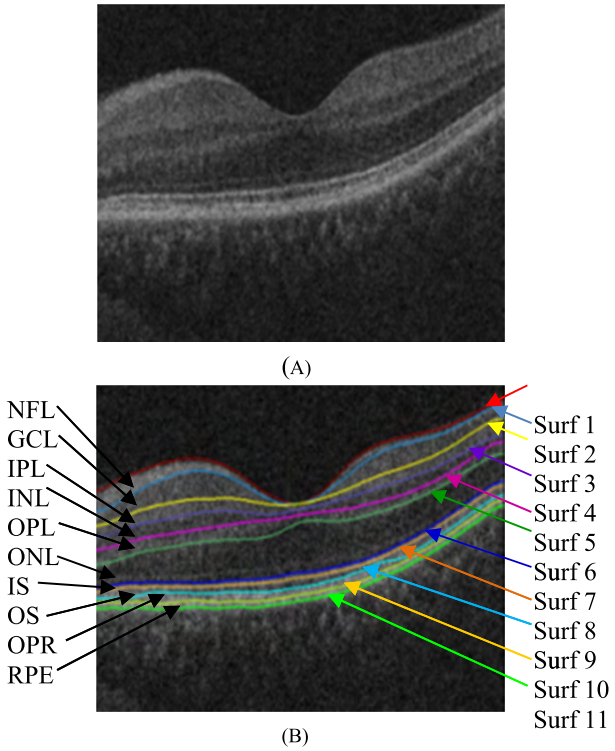


FIGURE 1. **(A) B-scan sample from normal eyes of an SD-OCT image. (B) 11 surfaces (Surf) from The Iowa Reference Algorithms [3] overlaid on the B-scan. The 11 surfaces delineate ten retinal layers, namely, the Nerve Fiber Layer (NFL), the Ganglion Cell Layer (GCL), the Inner Plexiform Layer (IPL), the Inner Nuclear Layer (INL), the Outer Plexiform Layer (OPL), the Outer Nuclear Layer (ONL), the Inner Segment (IS), and the Outer Segment (OS), Outer Photoreceptor (OPR), Retinal Pigment Epithelium (RPE).**

Despite their effectiveness in the segmentation of layers of normal eyes and layers with less topology deforming disease (such as drusen), the methods presented in [3]–[8], [10], and [11] do not achieve the same performance when the eyes are affected by a disease such as NRD. When NRD exists, the morphology of layers significantly changes and the continuity of a layer's intensity value may be highly interrupted. As a result, the segmentation of retinal layers is difficult.

For the segmentation of fluid associated abnormalities of the retina, a couple of references [13]–[21] have been presented. A 3D approach that uses probability constrained graph search and graph cut was presented [17]. The layer-dependent stratified sampling approach for intraretinal and subretinal fluid segmentation was presented in [16]. The serous pigment epithelial detachments (PED) segmentation framework that integrates the AdaBoost algorithm with multi-scale graph search and shape-constrained graph cut was presented in [18]. A recent study in [13] reported a NRD associated subretinal fluid segmentation method from SD-OCT images with significant accuracy.

The segmentation of layers and fluid regions has been addressed by some of the previous works. Zhang *et al.* [20] integrated 3D graph search and a supervised voxel classification method to segment the outer retinal-subretinal and fluid-filled abnormalities, respectively, from eyes with exudative age-related macular degeneration. Antony *et al.* [22] presented a method that segments microcystic macular edema and inner retinal surfaces using a graph-theoretic approach. Novosel *et al.* [15] presented an algorithm that uses the local attenuation coefficient contrast of layers surrounding the surface to segment the four surfaces between retinal layers and to delineate the boundaries of fluid-associated pathologies. Shi *et al.* [19] developed a method that segments 11 retinal surfaces and the PED volume from SD-OCT images. The 3D multi-scale graph search was applied to the layer segmentation. The intensity threshold and the height difference between the final two consecutive surfaces were used to detect PED. Roy *et al.* [23] used fully convolutional networks (FCN) to generate pixel-wise labeling for eight retinal layers and fluid regions.

However, to our knowledge, previous works do not address the segmentation of all discernible retinal layers from an SD-OCT image of NRD affected eyes. The proposal of a method that segments all noticeable retinal layers and NRD associated subretinal fluid is a very important contribution to clinical investigation. Compared to previous works, the innovative achievements of our work are as follows: (1) a novel method that learns constraints from normal eyes in order to segment 11 retinal surfaces, as well as the subretinal fluid from NRD affected eyes, was presented. (2) The "*divide and merge*" approach was introduced to reduce the computational cost of graph-based optimization. (3) The easy yet effective technique was introduced to locate the fovea region from the thickness map. (4) The combination of k-means cluster

and the rigid hard constraint was applied to estimate the interrupted and missed retinal layers.

II. METHODS

A. OVERVIEW OF THE PROPOSED METHOD

The proposed method consists of three phases including the prior information model; NRD associated subretinal fluid segmentation, and layer segmentation. Fig. 2 shows the flowchart of the proposed method. The prior information model calculates the mean and variance of the smoothness constraint from the layers of normal eyes segmented using The Iowa Reference Algorithms [3]. The graph cut technique was employed to segment the NRD associated subretinal fluid. Finally, in the layer segmentation phase, the graph search method was applied to segment the 11 retinal surfaces. The details of each phase were explained in the following sections.

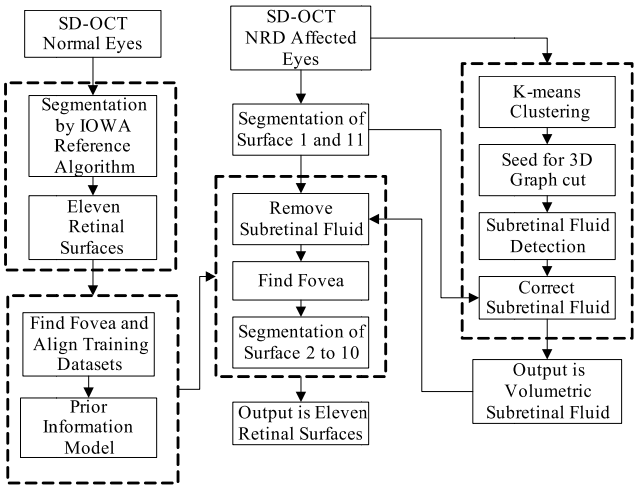


FIGURE 2. The flowchart of the proposed algorithm.

B. PRIOR INFORMATION MODEL

In this phase, the mean and variance of the distance smoothness constraint were calculated as the interval between a given surface and the reference surface. The purpose of a distance constraint is to limit the search area of the intended surface with respect to the reference surface. The constraint was learned from the fovea-centered 20 SD-OCT cubes of normal eyes segmented using The Iowa Reference Algorithms [3]. Since the left and right eyes have slightly dissimilar orientation, the training dataset for each eye was adopted from the eye with similar orientation. Alternatively, inverting around the x-axis would be sufficient to compensate for the orientation difference for the left to right eye or vice versa.

The distance between the surfaces varies across the B-scan. For example, surface 1 has a curvature shape around the fovea region and the distance from the other surfaces is smaller in this region (Fig. 3(C). Nevertheless, most of the clinical SD-OCT scans are not perfectly centered in the fovea region. To acquire an accurate prior information model, we find the fovea center and align the training datasets by orienting the fovea as the origin.

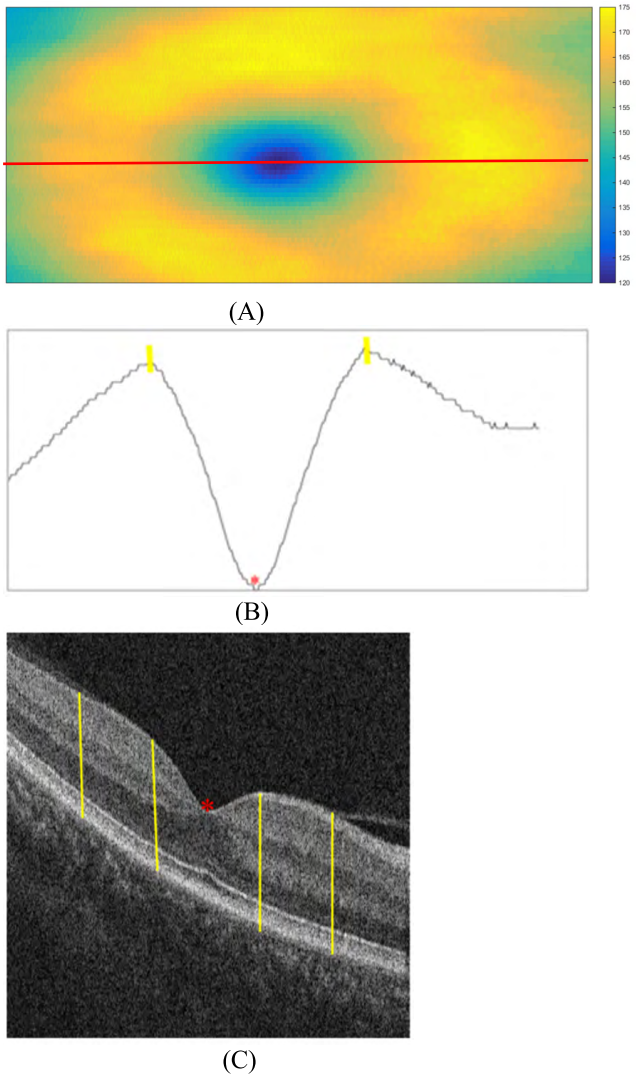


FIGURE 3. (A) Sample thickness map from a training dataset. The red line shows the row where the fovea center was located. (B) The graph representation of a thickness map row where the fovea center was located. The fovea center and maximum peaks are shown using the red asterisk and yellow lines, respectively. (C) The B-scan overlaid with the fovea center and five slices including the fovea region. The vertical yellow lines correspond to the boundaries where slices were partitioned. Note that the two middle yellow lines and the red asterisk are the peaks and the fovea center shown in panel (B).

The fovea center was estimated from the thickness map of the retina obtained for each training dataset by calculating the height difference of surfaces 1 and 11 Fig. 3(A). As a retina has its smallest thickness in the fovea center, the A-scan with minimum thickness was chosen as the fovea center. For the correct approximation, some B-scans (first and last 30 B-scans) and A-scans (100 left and right A-scans) were not considered.

The mean and variance of the distance constraint were calculated per slice of the B-scan. In this context, a slice is defined as a portion of a B-scan where a given surface and the referenced surface have a relatively similar height difference Fig. 3(C). Excluding the fovea region, the B-scan can be

divided into a random number of slices, but it is important to find the full fovea region and treat it separately. The full fovea region was bounded from the fovea center within a specified radius. The radius was assimilated by tracing two maximum peaks in a B-scan where the fovea center is located Figs. 3(A) & (B). This approach enables to find different sized fovea regions without estimation of a fixed radius. The Savitzky-Golay filter [28] was applied to the thickness map to decrease the effect of small peaks.

For surfaces S_i and S_j from n training datasets and B-scan k, the mean distance (μ_h) of slice h with the width of m was found using (1). The variance was also calculated accordingly.

$$\mu_h^{i,j} = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{m}\sum_{l=1}^{m}\left|S_i^k(x_l, y_l) - S_j^k(x_l, y_l)\right| \quad (1)$$

C. NRD ASSOCIATED FLUID SEGMENTATION

The segmentation of NRD associated subretinal fluid was conducted using a 3D graph-cut technique. A graph cut technique segments a region approximately with uniform intensity values by integrating the regional and neighborhood information. In this paper, we extended the interactive graph cut algorithm presented by Boykov and Jolly [29] into a fully automated approach. The details can be found in [29], but for consistency here we summarized the main idea.

For the segmentation of a given region, it is normal to consider the nature of the region and the property of its boundary. According to [29], the region segmentation can be formulated by using the energy minimization function, as follows:

$$E(f) = \sum_{v\in V} R_v(f_v) + \sum_{v\in V, u\in N_v} B_{vu}(f_v, f_u) \quad (2)$$

where R_v (f_v) and B_{vu} (f_v, f_u) are terms to describe the region and boundary properties, respectively, and N_v is the set of the neighborhood voxels of v.

The region term $R_v(f_v)$ evaluates the similarity of a voxel to be labeled an object or background, and penalizes for assigning a label f_v to voxel v. The boundary term $B_{vu}(f_v, f_u)$ explains the continuity of the neighborhood pixels and penalizes for labeling a pair of voxel v and u with the labels of f_v and f_u, respectively.

For the 3D graph cut, it is important to create a known model for the voxels' evaluation. For example, during binary segmentation such as object and background, the sample voxels from each label could be indicated as seeds, either automatically or by the user. Then, during the evaluation, if the voxel of the region is similar (e.g., in their intensity) to the sample voxels of the label, then the value of the region term is large and it is close to zero when they are different [29].

To solve the problem using a s/t cut, the graph was constructed as follows. The weight of the t-links that connect each node to both s and t terminals was calculated as the negative log-likelihood of the histogram of the intensity values of the seed voxels.

$$R_v(obj) = -lnPr\left(V_{inten} \mid obj_{seed}\right) \quad (3)$$

$$R_v(bkg) = -lnPr\left(V_{inten} \mid bkg_{seed}\right) \quad (4)$$

For the n-links that connect neighborhood voxels, the boundary penalty was calculated using (5), in which I_v and I_u are the intensity values of the neighbor pixels v and u. Equation (5) assigns a large weight when the neighbor voxels have similar intensity values and less weight when the pixels have different intensity values. Hence, the lowest value assigned for the boundary pixels that can be segmented using the s/t cut method as presented in [30] and implementation in [31].

$$B_{(v,u)} = \exp\left(-\frac{(I_v - I_u)^2}{2\sigma^2}\right).\frac{1}{dist(v,u)} \quad (5)$$

In the original work [29], the seeds were provided manually by the user, but in our method, they were found automatically by applying a k-means algorithm to the individual B-scans. The k-means algorithm is a popular algorithm that groups pixels of an image into k number of clusters. In the k-means cluster, it is important to determine the number of clusters k. To define the number of clusters, the three major reflectivity profiles of a retina were considered. Each B-scan displays the low reflectivity (vitreous fluid, sclera, and subretinal fluid region for the NRD affected), medium reflectivity (plexiform and the nuclear layers) and the high reflectivity (the nerve fibers and the retinal pigment epithelium) regions Fig. 4(B). As the k-means algorithm considers only the intensity values of pixels, it is important to utilize the 3D spatial connectivity information of voxels to obtain true subretinal fluid regions.

During the acquisition of an SD-OCT image, a very dark sclera region may be included and it can result in incorrect k-means clusters. To remove the possible incorrect clustering, the top and bottom rows (100 rows) of each B-scan were excluded.

The seeds for fluid (object) and retinal region (background) were based on the threshold value from the low reflectivity region. The threshold value can be the maximum value from the low reflective region or determined by substituting the mean and variance of the low reflectivity region into (6) with various standard deviations (see section III (A) for more details). The pixels with an intensity value less than the threshold were seeds for the object, whereas those with higher values were used for the background. During subretinal fluid segmentation, all the layers were considered to be background.

$$T = \mu_{lr} + \alpha_0.\sqrt{\sigma}_{lr} \quad (6)$$

where T stand for the threshold value. μ_{lr} and σ_{lr} are mean and variance from the low reflective region. α_0 stands for the standard deviation intervals that assumes the intensity distribution in the low reflective region is normally distributed.

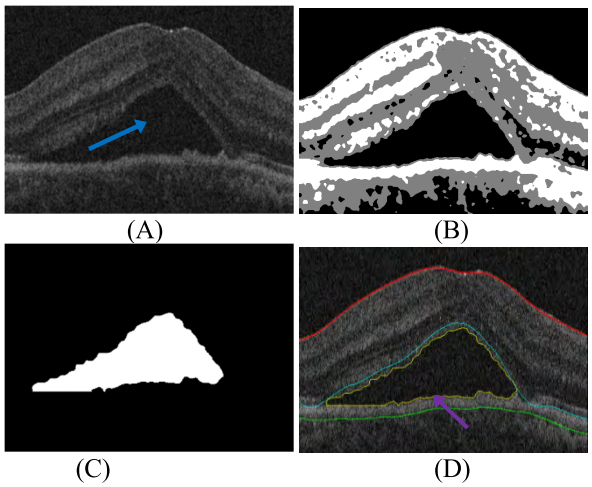


FIGURE 4. NRD associated fluid segmentation. (A) The B-scan shows the retinal region affected by NRD (the blue arrow indicates NRD related subretinal fluid). (B) The result from k-means clustering. Low, medium and high reflectivity regions were indicated using black, gray and white colors, respectively. (C) The binary image of detected NRD associated subretinal fluid. (D) B-scan overlaid with the results of subretinal fluid, surface 1, 9 and 11 (outlined in yellow, red, cyan and green colors respectively). The purple arrow shows the lower boundary of the fluid.

In this segmentation, because of the reduced intraretinal reflectivity, the outer nuclear layer (ONL) might be segmented as fluid. To remove the ONL region, the distance between the lower boundary of the segmented fluid region (Ex. the purple arrow in Fig. 4(D)) and surface 11 was checked. If the distance is greater than the minimum possible distance between surfaces 11 and 9 based on the prior information model, then it was removed. In other words, the fluid regions with the lower boundary above surface 9 (cyan color in Fig. 4(D)) were omitted. Since the NRD associated subretinal fluid appears below surface 9, this technique successfully removes the possible ONL segmentation. The maximum distance was calculated from the prior information model, preventing the user from having to estimate a threshold for different sized subretinal fluid. We applied the level set [32], [33] to smooth the boundary of the final result (Fig. 4(C) & (D)).

D. LAYER SEGMENTATION

Assume that X, Y, and Z denote the sizes of the SD-OCT images in the x, y, and z directions, respectively (Fig. 5.(A)). The boundary between layers was modeled as a terrain-like surface that intersects each column at exactly one voxel. This surface can be defined by a function $\rightarrow$ f : (x, y) z. To ensure an accurate representation of the surface, the shape information can be enforced.

The graph search method was used to segment the retinal surfaces. To apply the graph search, finding the optimal surface was transformed into the minimum closure problem solved by the minimum s-t cut [34]. The weighted graph $G = (V, E)$ with a set of vertexes V, and directed edges E of non-negative weights was constructed as shown in Fig. 5(B).

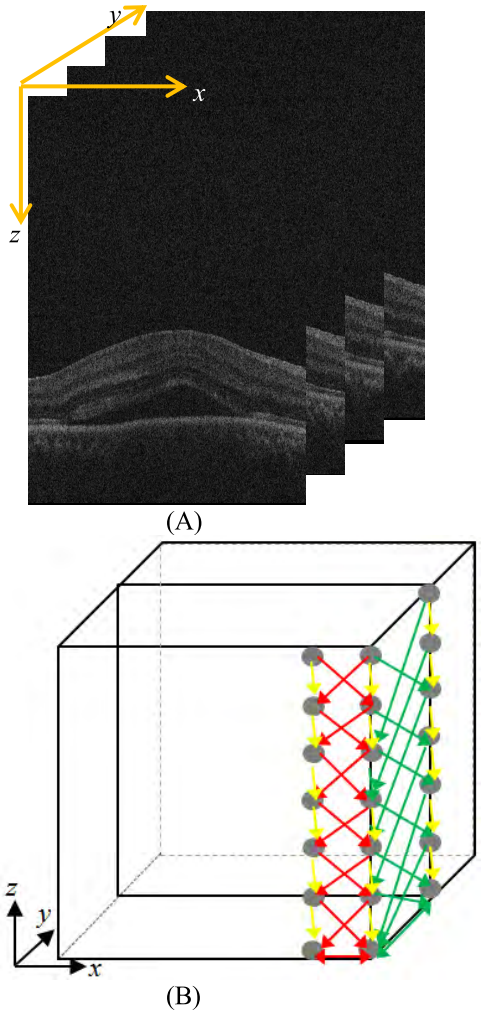


FIGURE 5. Sample SD-OCT cube and the graph construction illustration. (A) The orange arrows indicate the x, y, z directions. (B) The intra-column edges are shown by yellow arrow whereas the inter-column edges are shown using red (x-direction with a value of 1 for Δx) and green arrows (y-direction with a value of 2 for Δy).

The voxels of images represent the vertexes of the graph and the edges were defined as follows. The intra-column (yellow arrows) edges between a vertex $V(x, y, z)$ and $V(x, y, z-1)$ were created to ensure the intersection of the surface with each column exactly once. Given two columns (x, y) and (x', y) adjacent in the x-direction, the inter-column (red arrows) edges were added from $V(x, y, z)$ to $V(x', y, z - \Delta x_{min})$ and $V(x', y, z + \Delta x_{max})$. Δx_{min} and Δx_{max} are the minimum and maximum height changes of the adjacent columns in the x-direction. The same method was used in the y-direction (green arrows). The inter-column edges were intended to enforce the shape information of the surfaces.

The cost function $C(x, y, z)$ of each vertex was calculated as the inverse likelihood of the surface and its weight, formulated as Eq. (7). Each vertex was connected to the additional s and t nodes. The s node was connected to every vertex through a t_link with a weight of w_v and every vertex was connected to a t node with a weight of $-w_v$. The intra-column

and inter-column edges are commonly called the $n\_$ links and were assigned infinite weight.

$$w_v(x,y,z) = \begin{cases} c_v(x,y,z) & if\ z = 0 \\ c_v(x,y,z) - c_v(x,y,z-1), & otherwise \end{cases} \tag{7}$$

The gradient function that enhances the boundary of layers based on the transition from dark-to-bright or bright-to-dark was used to define the cost function. From top to bottom, surfaces 2, 4, 6, 8 and 10 have a bright-to-dark transition, whereas surfaces 1, 3, 5, 7 and 9 have a dark-to-bright transition. The anisotropic filter [35] was applied to each B-scan to reduce the noise, and then a Sobel operator was used to calculate the gradient.

In the case of multiple surface segmentations, we consider the distance between the surfaces to improve the segmentation result. For instance, if the surfaces are non-crossing, we can assume that a given surface is always below or above another surface. Thus, the distance between the surfaces can be incorporated by introducing additional smoothness constraint Δd, which imposes a minimum (d_{min}) and maximum (d_{max}) distance between the surfaces (8). If the minimum and maximum distance between a given surface and its reference surface are known, the search area of the surface can be limited to be within that range.

$$\Delta d_{min} \leq v_i(x,y,z) - v_j(x,y,z) \leq \Delta d_{max} \tag{8}$$

where $v_i(x,y,z)$ and $v_j(x,y,z)$ are voxels on the surfaces of S_i and S_j, respectively. and $\Delta d_{max} > \Delta d_{min} > 0$.

In our method, the distance smoothness constraint was learned in the prior information model phase. The mean and variance of the constraint from the prior information model enable us to determine the minimum and maximum of the constraints using Eq. (9). The distance information was applied to limit the search area of the surface with respect to the reference surface.

$$\begin{aligned} \delta_{min}(x,y) &= \mu_{(x,y)} - \alpha_1 . \sqrt{\sigma}_{(x,y)} \\ \delta_{max}(x,y) &= \mu_{(x,y)} + \alpha_2 . \sqrt{\sigma}_{(x,y)} \end{aligned} \tag{9}$$

where $\delta_{min}(x,y)$ and $\delta_{max}(x,y)$ represent the upper and lower boundary of the search area, $\mu_{(x,y)}$ and $\sigma_{(x,y)}$ are the mean and variance from the prior information model, and α_1 and α_1 are the standard deviation intervals. Assuming the prior information model is normally distributed, α_1 and α_2 represent the standard deviation intervals.

Including the prior information model, three approaches, the potential possible pixel extraction (for surfaces 1 and 11), the prior information model (for surfaces 2 through 6 and 8 through 10) and the combination of the two approaches (for surface 7) were employed to limit the search areas of the layers. Each approach was followed by a 3D graph search to obtain the final segmentation result.

The potential possible pixels of surfaces were located from binary images that were generated using threshold values from the k-means algorithm. For surface 1, the maximum intensity value of a low reflectivity cluster was used as the threshold, and from the resulting binary image, the first occurring non-zero pixels in the A-scan (Fig. 6(A)) were chosen as potential possible pixels. The potential possible pixels of surface 11 were estimated from the last occurring pixel on the binary image generated using the minimum value of high reflective cluster in a high reflectivity cluster. The existence of a higher intensity valued choroid (Fig. 6(B)) may cause incorrect potential possible pixel estimation for surface 11. To reduce the effect of the choroid, the morphological operators were applied to the binary images and rigid shape constraints were used in the 3D graph construction. To increase the search area, a fixed number of pixels (ten pixels) were allowed above and below the potential pixels and then the 3D graph search was applied to find the layers

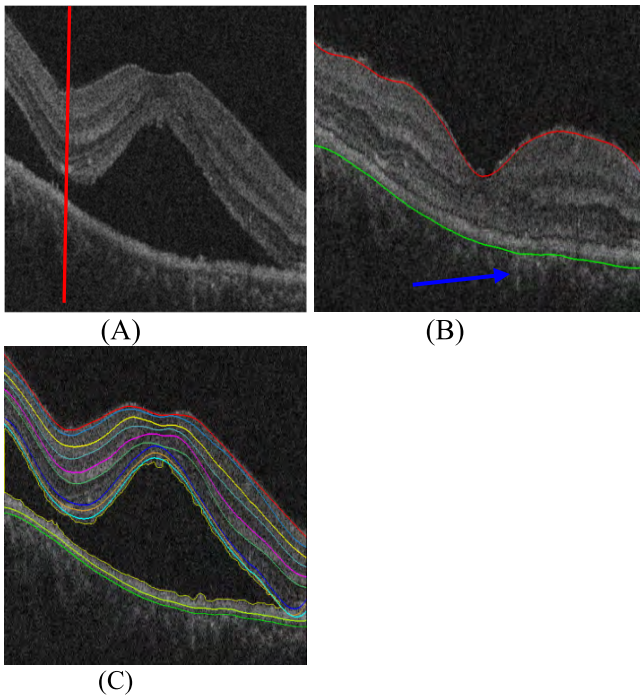


FIGURE 6. The eleven retinal surfaces from the NRD affected eye. (A) Original B-scan. The red line is an example of an A-scan (B) B-scan where the subretinal fluid was removed. The blow arrow shows the choroid region (C) The eleven retinal surfaces are overlaid on the original image.

The NRD associated subretinal fluid is known for changing the topology of layers. For instance, the curve nature of surface 1 around the fovea region may significantly alter when the subretinal fluid exist (Fig. 6(A)). For this reason, following the detection of the NRD related subretinal fluid, it was removed to attain the correct topology of the retina (Fig. 6(B)). In the A-scans where the subretinal fluid was removed, the height of surface 1 was adjusted. At this point, it is possible to locate the fovea center (as described in section II (B)) and align the image with the training datasets to adopt the prior information model.

In addition to the topological change, the distances between layers might deviate from the regular distance range,

which can affect the search area limitation using the constraint learned from normal eyes. To overcome this condition, we allow an extra search area by finding the upper and lower boundaries of the search area using the reference surface and the surface that is nearer and has a different transition than the intended surface. For example, surface 2 is under a bright-to-dark transition, whereas surface 3 is under a dark-to-bright transition. To determine the search area for surface 2, the lower boundary can be the maximum possible distance between surfaces 1 and 3, whereas the upper boundary is the minimum distance between surfaces 1 and 2. The maximum and minimum distance is calculated from the prior information model. The list of reference surfaces and chosen nearer surfaces are shown in Table 1.

The search areas of surface 7 utilize both the potential pixel finding approach and the prior information model. In other words, first, the search area was limited and then the occurrence of potential pixels was checked to guarantee the continuity of surfaces in each A-scan. The potential pixels, in this case, are the first occurrences of high reflectivity pixels in the bounded area; if these pixels are not found, they were searched for in the medium reflectivity cluster. Their absence in both clusters shows the pixels are missing, so we use interpolation to complete the missing values. This combination is important because in severe NRD, the intensity continuity of this layer is highly deteriorated and the gradient information was affected. Similar to surface 11, a fixed number of pixels were allowed above and below the potential pixels and the rigid constraint was applied in the 3D graph construction.

The accurate segmentation of the layers also depends on the shape smoothness constraints, which were modeled as the height changes of the surface between two consecutive A-scans both in the x-direction (Δx) and y-directions (Δy). In the x-direction, the surface shape changes smoothly when it was moved from one column to another. But in the y-direction, in addition to the shape changes, the eye movement artifacts can cause height changes. In the implementation of the proposed method, we have tested various combinations of shape smoothness constraints and the details were explained in the parameter estimation section. Finally, after the segmentation of all layers, the Savitzky-Golay filter [28] was applied to smooth the layers.

E. THE "DIVIDE AND MERGE" APPROACH

The main challenge in 3D graph-based approach is the computational time cost spent on graph optimization. In this work, we have used the *"divide and merge"* approach that decreases the computation time by reducing the graph size in the y-direction (Fig. 5). The diagrammatic representation is shown in Fig. 7. Upholding its order, the *"divide and merge"*approach splits the B-scans (the area between the dotted lines) of a cube into mini-batches (the area between red lines) to contain a limited number of B-scans for a given execution. The execution of the mini-batches was conducted by keeping the natural order of B-scans. After each execution, in order to maintain the topological continuity of the layers, the last few B-scans from preceding batch were merged to its succeeding batch (shaded area). That means, in addition to its B-scans, the execution of each succeeding mini-batch includes the B-scans from its preceding mini-batch. The dividing technique is to reduce the size of the graph whereas the merging preserves the topological continuity of layers.

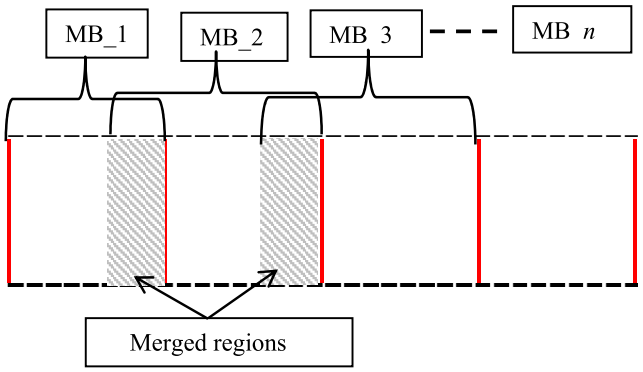


FIGURE 7. The diagrammatic representation of the "divide and merge" approach. MB stands for mini-batch. The area between the dotted lines represents the B-scans in the cube whereas the area between two red lines stands for the mini-batches for a given execution. The shaded regions are for the merged B-scans from consecutive mini-batches.

We have tested the *"divide and merge"*for different mini-batch and merging sizes and the result shows that accuracy of the result is not significantly affected by the batch based implementation (see the result in section III (E)). In our assumption, the reason for less effect of the batch system on accuracy is the structural anatomy of the retina which changes smoothly across the B-scan. The smooth shape enables the graph constructed from local B-scan to represent the structures of the layers.

III. EXPERIMENTS AND RESULT ANALYSIS

The proposed method was tested on 20 SD-OCT cubes of eyes diagnosed with NRD. The SD-OCT cubes were obtained from a Cirrus SD-OCT machine (Carl Zeiss Meditec, Inc., Dublin, CA). Each SD-OCT cube has 128 B-scans of 512 × 1024 pixels (512 A-scans where each A-scan comprises 1024 pixels).

To evaluate the proposed method, for the surfaces between the layers and the boundaries of NRD associated subretinal fluid, two experts created manual results. Surfaces 1, 2, 4, 5, 6, 7 and 11 were traced. The remainders of the surfaces were excluded because in the most severely affected retinal areas they are not visible to the human eye. The results of the algorithm, both surfaces and NRD associated subretinal fluid, were compared with the manual results from the experts. The accuracy of the results was further compared with published works. The surface segmentation was compared with the results from a fully convolutional network (FCN) [36] and '*DeepLabv3*' [27] whereas the subretinal fluid segmentation is in accordance with the recently published work by Wu *et al.* [13].

The precision of the surfaces' locations was examined by the metrics of the correlation coefficient *(cc)*, *p*-value and absolute surface position difference *(ASPD)*. For the segmentation results of the same B-scan from the manual segmentation result (Y) and the proposed method (X), the absolute surface position difference measures the mean difference in the axial direction (i) for each A-scan was calculated using (10). K represents the number of B-scans used for the evaluation and n is the number of columns across the B-scan. The standard deviation calculated in the usual way.

$$\overline{ASPD}\,(x, y) = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{n}\sum_{i=1}^{n}\left|X_k^i - Y_k^i\right| \qquad (10)$$

The accuracy of the NRD volume segmentation was calculated in terms of Dice-coefficient *(DC)*, overlap ratio *(Ovrlap)*, overestimated ratio *(Ovrest)* and underestimated ratio *(Undest)*. [37], [38]. The overlap ratio indicates the amount of NRD region available both in the results of the algorithm and of manual segmentation.

A. PARAMETER SELECTION

The parameters were fine-tuned on additional 13 manually segmented validation datasets.

The seeds for the subretinal fluid segmentation were obtained using threshold values estimated by (6), with the mean and variance of the low reflectivity region from the k-means cluster. Although the NRD is shown as a well-demarcated fluid region in SD-OCT images, severe cases commonly weaken the intensity distribution of the ONL and lead to miss cluster of this region under a low reflectivity cluster. Depending on the level of ONL interruption, various standard deviation intervals (α_0) can be replaced to estimate the threshold in (6). Assuming that the low reflectivity k-means region is normally distributed, a lower value of standard deviation interval can be substituted to promote uniform region segmentation and discourage the inclusion of a weak ONL as a fluid region. For a stronger boundary, the maximum value from a low reflectivity cluster or a threshold using a higher standard deviation interval can be substituted.

On the validation datasets, we have used the standard deviation of 2.5 for two cubes where ONL is strongly deteriorated and for the remainder of the testing datasets the maximum value of the low reflectivity cluster was used. For the reported results in section III (C & D), the maximum value of the low reflectivity cluster was used.

The estimation of the upper and lower boundaries of the layer search area uses (9). The standard deviation intervals of 2.7 for the lower boundary (α_1) and 2.9 for the upper boundary (α_2) were used so that 99 % of the possible search area was included. This enables us to locate the surfaces even when the thickness of the layers largely deviates from normal eyes because of the NRD related subretinal fluid. The list of reference surfaces and chosen nearer surfaces are shown in Table 1.

To determine the shape smoothness constraints, the shape variation of the layers and the strength of the gradient information were considered. Since in the x-direction the shapes of the surfaces change smoothly, Δx is set to 2 for all surfaces except surface 7 and 11 which were set to the value of 1.

TABLE 1. Details of shape constraint, reference surface and chosen nearer surface for search area limitation.

Segmentation Order	Surface No.	Reference Surface No.	Nearer Surface No.	Δy
1	1	No	No	8
2	11	No	No	1
3	2	1	3	5
4	4	2	5	4
5	3	2	4	4
6	7	11	5	1
7	6	2	4	4
8	5	1	3	4
9	9	7	No	2
10	8	7	10	2
11	10	11	No	2

The effect of Δy value selection on the absolute surface position difference was tested by setting its values from one to ten. As shown in Fig. 8, for surfaces 1, 2, and 6 the errors are smaller as the value of Δy increases and become stable for further increments of the values. This depicts that higher values of Δy can be tolerated by layers with higher gradient information. Surface 11 showed an increase in the error as Δy is larger because of the false positive potential pixels from the high reflectivity choroid region during the k-means cluster. The values of Δx and Δy are set to one to reduce the effect of false positive potential pixels. Due to intense value degradation in the severely affected retinal regions, surface 7 exhibited stability for small to medium values of Δy, but the highest error for larger values of Δy. Surfaces 4 and 5 are relatively stable in all ranges of the values, but we have used a small value since the gradient strength of these layers is weak.

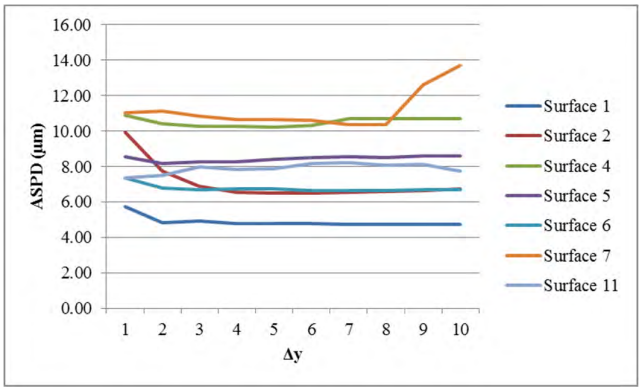


FIGURE 8. The absolute surface position difference (ASPD) of retinal surfaces with respect to various Δy values.

The surfaces prone to intensity value interruption (surfaces 8 through 10) have small values of the shape smoothness constraint, as shown in Table 1. By reducing the flexibility of the graph search, a small value of Δy controls

the possible wrong segmentation because of a disrupted intensity value. The surfaces 7 and 11 have Δy of 1 which impose rigid graph construction. Note that these results were obtained by changing the values of the intended surface while maintaining the values of the reference and nearer surfaces at the values listed in Table 1. Henceforth, the error propagation between layers was not considered.

B. CONVOLUTIONAL NEURAL NETWORK (CNN)

We have compared the proposed method with two CNN models named FCN [36] and '*DeepLabv3*' [27]. As described in section II (D), for the proposed method, the constraints were learned from normal eyes and the retinal surface segmentation was conducted after the detection and removal of the subretinal fluid region. For a fair comparison, we followed the same rule for CNN architectures such that they were trained using the retinal layers of normal eyes obtained from The Iowa Reference Algorithms [3] and tested on subretinal fluid-free B-scans in which the subretinal fluid was first detected and removed by the proposed method. During the evaluation phase (section III (C & D)) the fluid region was retained to attain the topology of the retina that matches the manual segmentation result. It is important to note that for the proposed method, the prior information was calculated as the distances between eleven retinal surfaces segmented using [3], whereas the training sets of CNN architectures are the ten retinal layers from [3], (see Fig. 1(B)).

CNN involves a number of hyper-parameters which requires careful selection to achieve an optimal solution. In practice, for different CNN architectures and datasets, the subset of hyper-parameters that matter most are different, but it is often the case that a few hyper-parameters made a significant difference [39]. In our experiment, we have fine-tuned the hyper-parameters listed in Table 2 using the random grid-based approach. The rest of the hyper-parameters are the default values suggested by the authors. The CNN architectures run on NVIDIA GeForceGTX 1080 GPU

TABLE 2. The list of hyper-parameters and chosen values.

Hyper-parameters	FCN	*'DeepLabv3'*
Learning rate	$1\times e^{-5}$	$1\times e^{-7}$
Batch	10	6
dropout probability	50%	
Number of learning iteration	20,000	25,000
Stride size		16

The codes for the CNN architectures were implemented using Tensorflow and downloaded from [40] and [41] for [36] and [27], respectively. We have used 2050 images for training and 500 for validation. The images were down-sampled by 2 to get the size of 512×512.

In both architectures, after 30,000 training iterations the overfitting started by increasing the validation error while keep decreasing the training error. To overcome the overfitting issue, we have used an early stop approach and constrained the training iteration number for the values mention in Table 2. For both [36] and [27] the training and validation accuracy curves with respect to the number of the iteration are shown in Fig. 9.

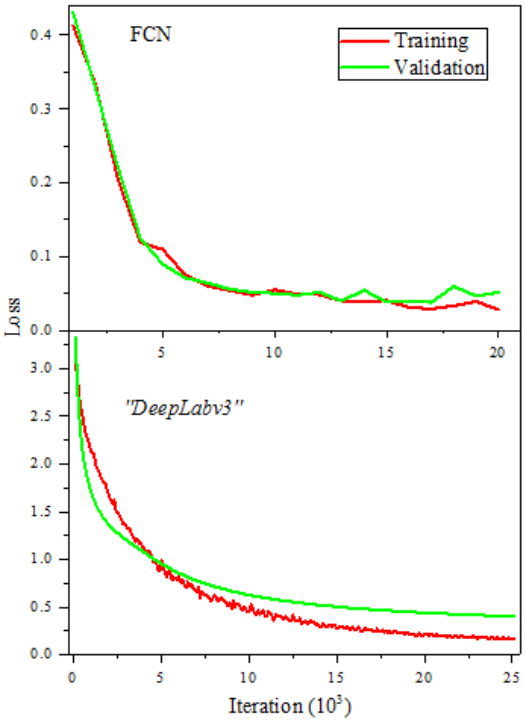


FIGURE 9. The training and validation loss of the CNN architectures.

C. QUALITATIVE AND QUANTITATIVE EVALUATION OF THE PROPOSED METHOD WITH CNN

The qualitative evaluation of the surface segmentation result of the proposed method was presented in contrast to [36] and [27].

Fig. 10 shows the example of the qualitative comparison of the proposed method with [36] and [27]. The proposed method provided accurate surface segmentations in contrast to [36] and [27] for most of the surfaces. Although the topology of the retina was retained by removing the subretinal fluid, the layers' thicknesses change and the intensity value interruption of the NRD affected region remain, which could be the reason for the incorrect surface segmentation using [36] and [27]. The results from CNN architectures experience a discontinuity in the cases of NRD associated fluid region (Figs. 10(C) & (D)). The CNN architectures segment surface 1 relatively well because the NFL layer was less interrupted by the NRD related subretinal fluid. The proposed approach is capable of tolerating structural as well as intensity value interruption of NRD affected regions.

The qualitative comparison of the subretinal fluid from the proposed method with [13] is shown in Fig. 11. The algorithm

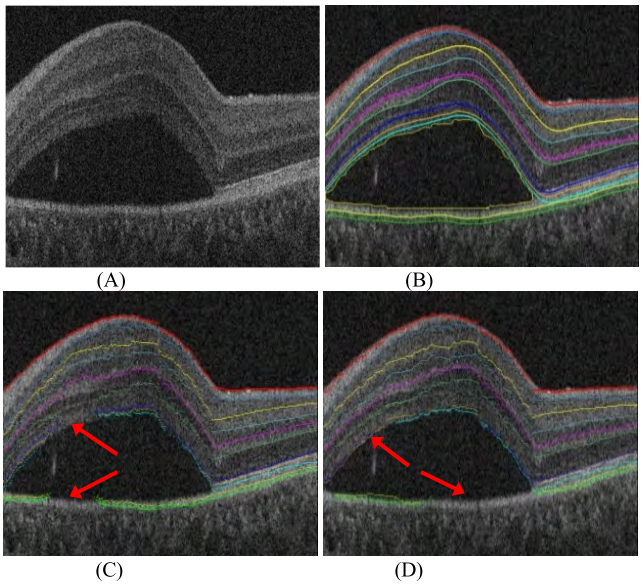


FIGURE 10. Examples of surface segmentation results from the proposed method, [36] and [27]. (A) the original B-scans, (B) the results of the proposed method, (C) and (D) are the results of [36] and [27] respectively. The red arrows on the last column show the discontinuity of surfaces.

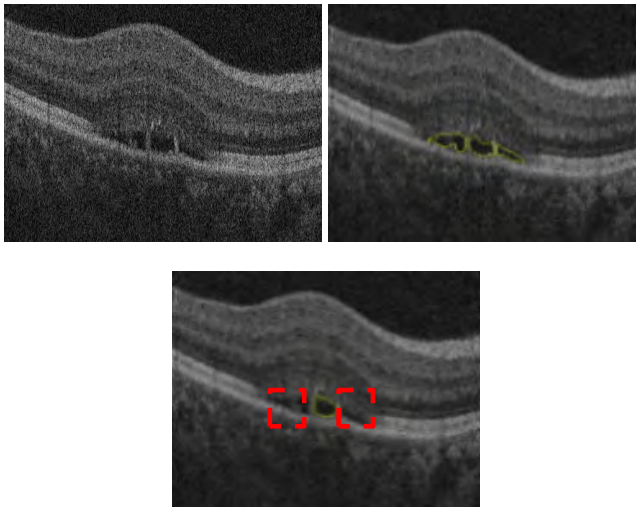


FIGURE 11. Examples of NRD associated subretinal fluid segmentation results. From left to right: the original B-scans, the results of the proposed method and [13].

in [13] failed to segment small subretinal regions (the red dashed boxes), whereas the proposed method successfully segmented the small regions.

D. QUALITATIVE EVALUATION

Table 3 shows the mean and standard deviation of the absolute surface position difference (ASPD) between the results of the proposed method (Alg.) and CNN with the baseline from the manual segmentation of the first expert (Exp. 1), and the inter-experts variability (Exp. 1 and Exp. 2). Compared to CNN, the proposed method depicted smaller ASPD for all the surfaces. The overall absolute position difference of the proposed method is $6.34\pm2.6\mu$ m with the first expert and the inter-expert variation is $6.39\pm5.9\mu$m, which is statistically comparable with the overall absolute position difference of the proposed method. The overall ASPD of [36] is 50.72 ± 57.75 μm and for [27] it is 92.24 ± 57.5 μm, which is larger than both the proposed method and the inter-expert variability. *'DeepLabv3'* [27] showed a larger deviation from manual results compare to FCN [36].

Compare to the values presented in Table 3, the result in Fig. 8 may show a visually better result. Yet it is important to mention that Fig. 8 was anticipated to demonstrate the discontinuity of only surfaces 7-9. In other testing images, the rest of the layers were also not detected properly which is the reason for the higher value of ASPD in Table 3.

Table 4 presents the evaluation of the retinal surfaces variability in terms of p-value and correlation coefficient (cc). The higher p-value in the U-test indicates that the surfaces from the methods do not have a statistically significant difference. The correlation coefficient of the proposed method is higher and comparable with that of the inter-expert variability. FCN [36] showed a higher p-value and correlation coefficient for surface 1, and has significantly smaller values for the other surfaces whereas "*DeepLab*" [27] showed less co-relation with manual results. The p-value ($p < 0.05$) in the U-test indicates that there is a statistically significant difference compared to the baseline.

The quantitative evaluation of the absolute surface position difference in Table 3 and the p-values with the correlation coefficients in Table 4 show that the proposed method has higher accuracy than [27] and [36], and it is comparable with the inter-expert variability. Therefore, the proposed method is sufficiently accurate to segment layers from NRD affected eyes.

The result of the quantitative evaluation of NRD associated subretinal fluid is shown in Table 5. The results from the proposed method (Alg.) and those obtained by running the algorithm presented by [13] were compared to the manual segmentation results from two experts. The overlap and dice coefficient shows that the results from the proposed method showed statistically significant overlap with the results from the experts. The overlap and underestimate ratio of our method with the experts is comparable to that of [13]. Our method showed less overestimate compared to [13], which shows that the proposed method segments less false positive NRD regions compared to [13].

E. EVALUATION OF THE "DIVIDE AND MERGE" APPROACH

The proposed method was implemented using Matlab R2014b and C++, and run on an Intel i5-3317 CPU @ 1.70 GHz PC with 4 GB memory. To achieve efficient computational time, the B-scans were down-sampled by two.

The computational time (Cm. time in seconds) and absolute surface position difference (ASPD in μm) for different combinations of mini-batches (16, 32 and 64) and merge size (2 and 5) were shown in Table 6.

TABLE 3. Absolute surface position difference for NRD affected data.

Surfaces	Alg. vs. Exp. 1	Exp.1 vs. Exp. 2	[36] vs. Exp. 1	[27] vs. Exp. 1
1	4.18±1.33	5.08±4.26	8.03±15.04	11.24±16.42
2	4.33±1.78	6.08±4.83	7.43±4.68	12.56±13.76
4	7.05±3.4	7.17±6.16	12.63±13.35	18.18±9.99
5	6.84±3.51	6.06±6.18	13.19±17.46	20.96±15
6	4.96±3.07	6.99±6.65	94.56±117.35	64.96±94.6
7	8.36±3.46	6.38±8.29	112.33±121.49	254.24±100.77
11	8.62±1.74	6.96±4.91	106.85±114.89	263.53±151.92
Overall	6.34±2.6	6.39±5.9	50.72±57.75	92.24±57.5

Note: 1 pixel = 1.95 µm

TABLE 4. *P*-value and correlation coefficient for NRD affected data.

Surfaces	Alg. vs. Exp. 1		Exp.1 vs. Exp. 2		[36] vs. Exp. 1		[27] vs. Exp. 1	
	p-value	*cc*	*p*-value	*cc*	*p*-value	*cc*	*p*-value	*cc*
1	0.4678	0.9994	0.4672	0.9993	0.4583	0.9512	0.1716	0.9512
2	0.5661	0.9991	0.4601	0.9989	0.3686	0.9879	0.2575	0.9578
4	0.4081	0.9974	0.4881	0.9986	0.2628	0.9757	0.0503	0.9838
5	0.3260	0.9990	0.4064	0.9995	0.2778	0.9871	0.0023	0.9797
6	0.4838	0.9989	0.3801	0.9993	0.1916	0.6811	0.0590	0.7838
7	0.3109	0.9962	0.4704	0.9979	0.1711	0.5512	0.1118	0.5190
11	0.2836	0.9985	0.3986	0.9988	0.0117	0.4314	0.0020	0.2143

TABLE 5. Volumetric segmentation result of NRD associated subretinal fluid compared with manual and wu's method [13].

Terms (%)	Exp. 1 vs. Alg.	Exp. 2 vs. Alg.	Exp. 1 vs. [13]	Exp. 2 vs. [13]
Ovrlap	85.35±4.82	87.60±4.48	82.84±6.99	85.03±6.77
Ovrest	24.75±34.65	24.03±47.66	25.80±34.90	25.79±48.20
Undest	8.83±4.94	8.69±4.66	10.83±7.27	10.32±7.01
DC	90.78±14.49	92.04±13.98	87.40±22.62	88.69±22.30

TABLE 6. The computational time and absolute surface position difference for different combinations of minibatches and merge sizes (mean ± SD).

Merge size	Mini-batch size					
	16		32		64	
	ASPD	Cm. time	ASPD	Cm. time	ASPD	Cm. time
2	6.34±2.6	852.3±80.5	6.45±2.8	818.6±119.6	7.55±3.9	1666.5±262.9
5	6.4±2.7	857.4±80.0	6.48±2.7	1026.2±164.2	7.64±3.7	1901.4±222.5

For comparison purpose, the testing datasets were also executed without applying the *"divide and merge"* in which the ASPD was 8.4±4.08μm and the computational time was 3484.03±735.80s. The evaluation of the *"divide and merge"* approach shows that the computation time for the *"dive and merge"*is lower and the ASPD values have no significant difference.

IV. DISCUSSION AND CONCLUSIONS

We presented an automated 3D segmentation method for 11 retinal surfaces and subretinal fluid from SD-OCT images affected by NRD. The 3D graph search method was employed to segment the surfaces. The distance smoothness constraint for graph construction was found from the prior information model. The prior information model calculates the mean and variance of the constraint from retinal surfaces of normal eyes.

The NRD related subretinal fluid segmentation was conducted using a graph cut in which the seeds for the object and background were automatically found with a k-means clustering algorithm. The k-means algorithm partitions the B-scans into low, medium and high reflectivity clusters. The pixels with intensity values less than the maximum value of the low reflectivity cluster were used to find seeds for the

object and background. The results of the NRD associated subretinal fluid were compared with manual segmentation results and Wu's method [13]. The proposed method showed better achievement.

The main challenge for the layer segmentation of NRD affected eyes is the unpredictable change of the layers' topology and their intensity information interruption because of the subretinal fluid. The method showed that the detection and removal of fluid can be integrated with the classification result of the k-means cluster to retain the natural topology of the layers and estimate the layers' boundary where their intensity values were interrupted. We have introduced the *"divide and merge"* approach to reduce the computational time of graph optimization. This approach was evaluated and showed reasonable execution time improvement.

To evaluate the effectiveness of the approach we have trained two CNN models, FCN [36] and *'DeepLabv3'* [27] with layers from normal eyes and tested the subretinal fluid free B-scans from our proposed method. The proposed method utilizes the distance information from [3], yet for [27] and [36] the ten retinal layers were used for training. Compared to CNN models, the overall absolute surface position difference of the proposed method is statistically indistinguishable from the inter-observer variability.

Obviously, many research works [12], [23]–[26] have demonstrated the effectiveness of deep learning for retinal layer segmentation. The lower performance of the CNN in this test might be because of the used training datasets. For the tasks such as layer segmentation from NRD affected eyes, since layers experience structural changes, it is important to train the CNN with training dataset from NRD affected rather than the normal eyes. CNN would produce optimal solution if it was trained on a large number of accurate training datasets from NRD affected eyes that can represent different topological changes of layers for the model to extract feature that represents layers thickness changes and possible intensity value interruptions. For a larger number of training datasets, CNN provides better accuracy, yet because of the time-consuming nature of manual segmentation, finding training data for ten layers is very difficult.

For the discontinuities of layers as shown in Fig. 8, the CNN architectures also may produce a better result if post-processing like boundary tracing on probability map was applied.

Our proposed method reveals that graph search method can still be valuable for layers segmentation from NRD affected eyes with prior information that can be found using the publically available software as described in section II (B).

When the iteration number and the datasets were small the CNN models run on CPU. Though it is possible to run on CPU, on our training dataset we have seen that both FCN and *"DeepLabv3"*barely run on the computer used to run the proposed method (see the specification of the computer in see section III (E)). The FCN model iterated only 10 training iteration steps whereas *"DeepLabv3"* didn't run even single iteration in 12 hours. This is presumably because of the higher training time and resource demanding nature of the deep learning models. CNN requires higher training time compared to graph search and it is also computationally resource demanding because it needs high-performance graphics processing units and large amounts of storage. The proposed method can be used on any standalone computer.

The proposed method is sequential, in which subretinal fluid was segmented and then retinal surfaces segmentation follows. This architecture can lead to error propagation because incorrectly detected subretinal fluid can affect the retinal layer segmentation Fig. 12. Although the proposed method works well for subretinal fluid detection, the presence of blood vessels, which have a relatively similar intensity distribution as NRD associated subretinal fluid, can affect the results by increasing the overestimation term and resulting in an incorrect layer segmentation result. In future work, the proposed method can be extended to utilize the effective feature extraction ability of CNN to segment retinal layers from eyes affected by other subretinal fluids, such as the mostly co-occurring disease called serous pigment epithelial detachments (PED).

REFERENCES

[1] A. Daruich *et al.*, "Central serous chorioretinopathy: Recent findings and new physiopathology hypothesis," *Prog. Retin. Eye Res.*, vol. 48, pp. 82–118, Sep. 2015.

[2] M. Gemenetzi, G. De Salvo, and A. J. Lotery, "Central serous chorioretinopathy: An update on pathogenesis and treatment," *Eye*, vol. 24, no. 12, pp. 1743–1756, 2010.

[3] Iowa City, IA, USA, Iowa Inst. Biomed. Imag. *The Iowa Reference Algorithms*. Accessed: Mar. 25, 2018. [Online]. Available: https://www.iibi.uiowa.edu/downloads

[4] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," *Opt. Express*, vol. 18, no. 18, pp. 19413–19428, 2010.

[5] P. A. Dufour *et al.*, "Graph-based multi-surface segmentation of OCT data using trained hard and soft constraints," *IEEE Trans. Med. Imag.*, vol. 32, no. 3, pp. 531–543, Mar. 2013.

[6] M. K. Garvin, M. D. Abramoff, 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," *IEEE Trans. Med. Imag.*, vol. 27, no. 10, pp. 1495–1505, Oct. 2008.

[7] J. Novosel, G. Thepass, H. G. Lemij, J. F. de Boer, K. A. Vermeer, and L. J. van Vliet, "Loosely coupled level sets for simultaneous 3D retinal layer segmentation in optical coherence tomography," *Med. Image Anal.*, vol. 26, no. 1, pp. 146–158, Dec. 2015.

[8] 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," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1436–1447, Sep. 2009.

[9] T. Fabritius, S. Makita, M. Miura, R. Myllylä, and Y. Yasuno, "Automated segmentation of the macula by optical coherence tomography," *Opt. Express*, vol. 17, no. 18, pp. 15659–15669, 2009.

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

[11] A. Lang *et al.*, "Retinal layer segmentation of macular OCT images using boundary classification," *Biomed. Opt. Express*, vol. 4, no. 7, pp. 1133–1152, 2013.

[12] Y. He *et al.* (2018). "Topology guaranteed segmentation of the human retina from OCT using convolutional neural networks." [Online]. Available: https://arxiv.org/abs/1803.05120

[13] M. Wu *et al.*, "Automatic subretinal fluid segmentation of retinal SD-OCT images with neurosensory retinal detachment guided by enface fundus imaging," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 87–95, Jan. 2018.

[14] D. C. Fernandez, "Delineating fluid-filled region boundaries in optical coherence tomography images of the retina," *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 929–945, Aug. 2005.
[15] J. Novosel, Z. Wang, H. de Jong, M. van Velthoven, K. A. Vermeer, and L. J. van Vliet. "Locally-adaptive loosely-coupled level sets for retinal layer and fluid segmentation in subjects with central serous retinopathy," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 702–705.
[16] X. Xu, K. Lee, L. Zhang, M. Sonka, and M. D. Abràmoff, "Stratified sampling Voxel classification for segmentation of intraretinal and subretinal fluid in longitudinal clinical OCT data," *IEEE Trans. Med. Imag.*, vol. 34, no. 7, pp. 1616–1623, Jul. 2015.
[17] X. Chen, M. Niemeijer, L. Zhang, K. Lee, M. D. Abràmoff, and M. Sonka, "Three-dimensional segmentation of fluid-associated abnormalities in retinal OCT: Probability constrained graph-search-graph-cut," *IEEE Trans. Med. Imag.*, vol. 31, no. 8, pp. 1521–1531, Aug. 2012.
[18] Z. Sun *et al.*, "An automated framework for 3d serous pigment epithelium detachment segmentation in SD-OCT images," *Sci. Rep.*, vol. 6, Feb. 2016, Art. no. 21739.
[19] F. Shi *et al.*, "Automated 3-D retinal layer segmentation of macular optical coherence tomography images with serous pigment epithelial detachments," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, pp. 441–452, Feb. 2015.
[20] L. Zhang, M. Sonka, J. C. Folk, S. R. Russell, and M. D. Abràmoff, "Quantifying disrupted outer retinal-subretinal layer in SD-OCT images in choroidal neovascularization," *Invest. Ophthalmol. Vis. Sci.*, vol. 55, no. 4, pp. 2329–2335, 2014.
[21] G. Quellec, K. Lee, M. Dolejsi, M. K. Garvin, M. D. Abramoff, and M. Sonka, "Three-dimensional analysis of retinal layer texture: Identification of fluid-filled regions in SD-OCT of the macula," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1321–1330, Jun. 2010.
[22] B. J. Antony *et al.*, "Simultaneous segmentation of retinal surfaces and microcystic macular edema in SDOCT volumes," *Proc. SPIE, Med. Imag., Image Process.*, vol. 9784, p. 97841C, 2016.
[23] A. G. Roy *et al.*, "ReLayNet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Express*, vol. 8, no. 8, pp. 3627–3642, 2017.
[24] K. Gopinath and J. Sivaswamy, "Segmentation of retinal cysts from optical coherence tomography volumes via selective enhancement," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 1, pp. 273–282, Jan. 2019.
[25] F. G. Venhuizen *et al.*, "Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography," *Biomed. Opt. Express*, vol. 9, no. 4, pp. 1545–1569, 2018.
[26] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express*, vol. 8, no. 5, pp. 2732–2744, 2017.
[27] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam. (2017). "Rethinking atrous convolution for semantic image segmentation." [Online]. Available: https://arxiv.org/abs/1706.05587
[28] R. W. Schafer, "What is a Savitzky-Golay filter? [lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, Jul. 2011.
[29] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, Jul. 2001, pp. 105–112.
[30] A. V. Goldberg and R. E. Tarjan, "A new approach to the maximum-flow problem," *J. ACM*, vol. 35, no. 4, pp. 921–940, 1988.
[31] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.
[32] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.
[33] S. Niu, Q. Chen, L. de Sisternes, Z. Ji, Z. Zhou, and D. L. Rubin, "Robust noise region-based active contour model via local similarity factor for image segmentation," *Pattern Recognit.*, vol. 61, pp. 104–119, Jan. 2017.
[34] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 119–134, Jan. 2006.
[35] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[36] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[37] I. P. Okuwobi *et al.*, "Automated segmentation of hyperreflective foci in spectral domain optical coherence tomography with diabetic retinopathy," *J. Med. Imag.*, vol. 5, no. 1, p. 014002, 2018.
[38] Q. Chen, W. Fan, S. Niu, J. Shi, H. Shen, and S. Yuan, "Automated choroid segmentation based on gradual intensity distance in HD-OCT images," *Opt. Express*, vol. 23, no. 7, pp. 8974–8994, 2015.
[39] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade*. Springer, 2012, pp. 437–478.
[40] *The Tensorflow Implementation of 'Deeplabv3'*. Accessed: Jul. 15, 2018. [Online]. Available: https://github.com/sthalles/deeplab_v3
[41] *The Tensorflow Implementation of Fully Convolutional Networks for Semantic Segmentation*. Accessed: May 31, 2018. [Online]. Available: https://github.com/shekkizh/FCN.tensorflow

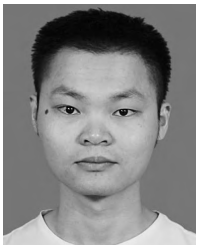
LOZA BEKALO received the B.S. degree in computer science from Hawassa University, Ethiopia, in 2009, and the M.Sc. degree in geoinformatics from the University of Twentee, The Netherlands, in 2014. She is currently pursuing the Ph.D. degree in computer science from the Nanjing University of Science and Technology, China. She was an Assistant Lecturer with the Debre Berhan and Hawassa Universities, Ethiopia. Her area of research interests is image analysis.

SIJIE NIU received the B.S. degree from the School of Computer Science, Liaocheng University, in 2007, and the Ph.D. degree from the Nanjing University of Science and Technology, in 2016. He was a Visiting Scholar with Stanford University. He is currently a Lecturer with the School of Information Science and Engineering, University of Jinan, China. His research interests include pattern recognition, machine learning, image analysis, and bio-information processing.

XIAOJUN HE received the B.S. degree in computer science from the Nanjing University of Science and Technology, in 2016, where she is currently pursuing the M.S. degree in computer science. Her main research contents are medical image processing and analysis.

PING LI received the B.S. degree in computer science from Jiangnan University, China, in 2015, and the M.Sc. degree in pattern recognition and intelligent systems from the Nanjing University of Science and Technology, China, in 2018. Her area of research interests is medical image processing.

IDOWU PAUL OKUWOBI received the B.S. degree in aeronautical engineering, and the M.S. degree in mechanical and electrical engineering from the Nanjing University of Aeronautics and Astronautics, China, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with the Nanjing University of Science and Technology, China. His current research interests include medical image segmentation and analysis, digital image processing, image segmentation, machine learning, and pattern recognition.



CHENCHEN YU received the B.S. degree in software engineering and the M.S. degree in pattern recognition and intelligent systems from the Nanjing University of Science and Technology, in 2015 and 2018, respectively. His research is about medical image processing and analysis.

WEN FAN received the M.D. degree from Wuhan University, Wuhan, China, in 2012. She is currently an Associate Chief Physician, and also an Assistant Professor with The First Affiliated Hospital, Nanjing Medical University, Nanjing, China. Her research interests include retinal imaging and vitreoretinal diseases.

SONGTAO YUAN received the M.D. degree from Tianjin Medical University, Tianjing, China, in 2006. He is currently an Associate Professor with Nanjing Medical University, Nanjing, China, and also a Senior Consultant of the Ophthalmology Department, The First Affiliated Hospital, Nanjing Medical University, Nanjing. His research interests include retinal imaging, analysis, and retinal diseases.

QIANG CHEN received the B.S. and Ph.D. degrees from the Nanjing University of Science and Technology, Nanjing, China, in 2002 and 2007, respectively. He held a postdoctoral position with the Stanford University, from 2010 to 2011. He is currently a Professor with the Nanjing University of Science and Technology, Nanjing, China. His research interests include image processing and analysis, and machine learning.