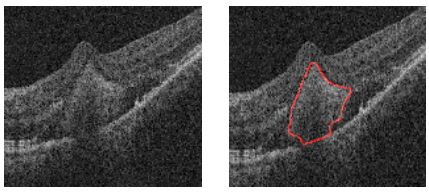
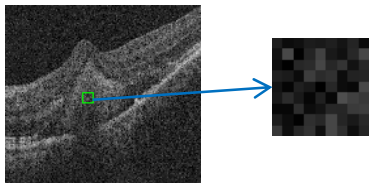
1. Challenge of CNV segmentation in OCT images



(a)Original Images (b)Ground Truth

Fig.1. Two slice image examples of CNV

Compared with other modality images, OCT images are more complicated in intensity distribution, which makes the existing methods failed to achieve satisfactory segmentation performance. Fig.1 shows two OCT image slices with CNV. We can observe that the CNV is a complicated object with irregular shape and size. Moreover, the intensity distribution is uneven within the CNV region. We also extract a local patch from CNV region, and calculate the intensity difference between central pixel and its neighboring pixels, as is shown in Fig.2. Fig.3 shows the intensity distribution of CNV and background. As shown in this figure, there is a large intensity overlap between CNV and background.



Original image Intensity variation image

Fig.2. A typical image with intensity variation. The intensity variation is generated by calculating the intensity difference between central pixel and other pixels in a 9× 9 neighbor region.

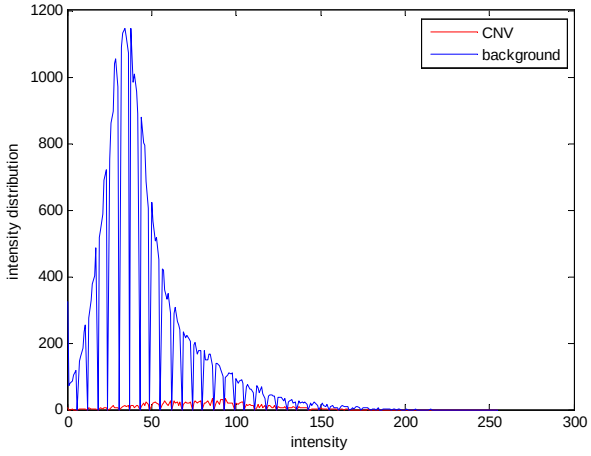


Fig.3. Intensity distribution between CNV and background

From these three figures, we can infer that the complicated intensity distribution may result in the large intensity variation between neighboring pixels within the same tissues. The uneven intensity distribution in local region may increase large intra-class variation, resulting in erroneous segmentation. Therefore, how to deal with the large intensity variation between neighboring pixels effectively is important to improve CNV segmentation performance.

2. Details of Experimental Results and Analysis

2.1 Experimental Setting

The experiments were performed on 15 OCT data with CNV, which were acquired using Topcon 3D-OCT-1000 (Topcon Corporation, Tokyo, Japan). Each SD-OCT volume contains 512×1024×128 voxels. This study was approved by the

Intuitional review board of Joint Shantou International Eye Center and adhered to the tenets of the Declaration of Helsinki. Because of its retrospective nature, informed consent was not required from subjects. The ground truth of CNV region in all B-scans is manually delineated by retinal specialists.

To evaluate the performance of the proposed method, Accuracy(ACC),Dice Similarity Coefficient (DSC), True Positive Volume Fraction (TPVF) and False Positive Volume Fraction (FPVF) were used as performance indices.

Accuracy indicates the segmentation accuracy of CNV pixels and background pixels. DSC was used to measure the accuracy of the automatic CNV segmentation result as compared against reference standard delineation; TPVF indicates the fraction of the true segmentation of CNV by the proposed method against the total amount of CNV; FPVF denotes the fraction of CNV pixels falsely identified by the proposed method. They are calculated as follows:

$$ACC = \frac{|V_A \cap V_M| + |(V - V_{A)} \cap (V - V_M)|}{|V|} \quad (18)$$

$$DSC = 2 \times \frac{|V_A \cap V_M|}{|V_A \cup V_M|} \quad (19)$$

$$TPVF = \frac{|V_A \cap V_M|}{|V_M|} \quad (20)$$

$$FPVF = \frac{|V_A| - |V_A \cap V_M|}{|V - V_M|} \quad (21)$$

Where $|\cdot|$ denotes volume, V_A denotes the CNV region segmented by the proposed method, V_M denotes the CNV region delineated by a retinal specialist, V denotes the total volume of the OCT data.

In our experiment, CNV detection result based on superpixels and SRC is used as the initial contour of MICO and ACM-LSP. In the Global intensity fitting energy, we set parameter $\boldsymbol{q}$ =2. In ACM-LSP, we set hyper parameter α=10.

2.2 Assessment of Segmentation Performance

This experiment shows the performance of proposed ACM-LSP model. In addition, considering the coarse position of CNV is detected by local similarity prior learning (LSPL) model, this experiment also gives the detection result of LSPL model.

Table 1 CNV segmentation result of ACM-LSP and LSPL

	ACM-LSP	LSPL	*p*-values
ACC	**0.9878±0.004**	0.9739+0.007	0.008
DSC	**0.5692±0.0484**	0.4677±0.070	0.037
TPVF	0.6239±0.1381	**0.8751±0.091**	0.005
FPVF	**0.0065±0.002**	0.0265±0.007	0.003

In this experiment, ACC,DSC, TPVF and FPVF of these two methods are listed in Table 1. Paired t-tests are used for the statistical test and a *p*-value less than 0.05 is considered statistically significant. For LSPL model, ACC, DSC, TPVF and FPVF are 0.9739, 0.4677, 0.8751, and 0.0265 while ACM-LSP are 0.9878,0.5692, 0.6239, and 0.0065 respectively. The *p*-values are 0.008,0.037, 0.005, and 0.003 respectively. As observed in Table 1, LSPL model achieves higher TPVF. However, ACM-LSP outperforms LSPL on ACC, DSC and FPVF.

According to the results, we can infer that coarse position of CNV can be detected accurately via LSPL model based on superpixel and SRC. The reason is that superpixel and SRC are robust to the intensity variation of pixels. By using the detection result of LSPL as the initial region, ACM-LSP boosts the performance significantly, especially for DSC, increasing about 10 percentage points. Based on the initial regions detected by LSPL model, ACM-LSP can capture more details by employing global intensity information and local similarity of pixels, leading to improvement of segmentation performance.

2.3 Effectiveness of local similarity prior evaluation

In this paper, the energy of MICO is used as the global intensity fitting energy of ACM-LSP. Therefore, we compare MICO and ACM-LSP to evaluate the effectiveness of local similarity prior in this experiment.

Table 2 CNV segmentation result, Compared with MICO

	ACM-LSP	MICO	*p*-values
ACC	**0.9878±0.004**	0.9813±0.008	0.008
DSC	**0.5692±0.0484**	0.3656±0.1317	0.037
TPVF	**0.6239±0.1381**	0.3343±0.0692	0.005
FPVF	0.0065±0.002	**0.0062±0.0035**	0.003

Table 2 lists the performance of ACM-LSP and MICO. As observed in this table, ACM-LSP outperforms MICO significantly, especially for DSC, increasing about 20 percentage points.

MICO enable deal with intensity inhomogeneity problem with the assumption that the bias field is varying slowly. However, it's difficult to achieve satisfied performance in our task because the intensity between the neighboring pixels may vary sharply.

As shown in Fig.4, for regions which have large intensity variation (the region in the blue bounding box), MICO only segments part of these regions due to the complicated intensity distribution. On the contrast, ACM-LSP can segment most of these regions accurately. Therefore, ACM-LSP outperforms MICO on segmentation of regions with large intensity variation. The reason is that ACM-LSP is developed with incorporation of learned local similarity prior. The learned prior map can capture the spatial information of CNV and spatial relationship between pixels, which can enforce the neighboring pixels within the same tissue have same label. Therefore, ACM-LSP is more robust to uneven intensity distribution in the local region, which can avoid the erroneous segmentation caused by large intensity variation between the neighboring pixels.

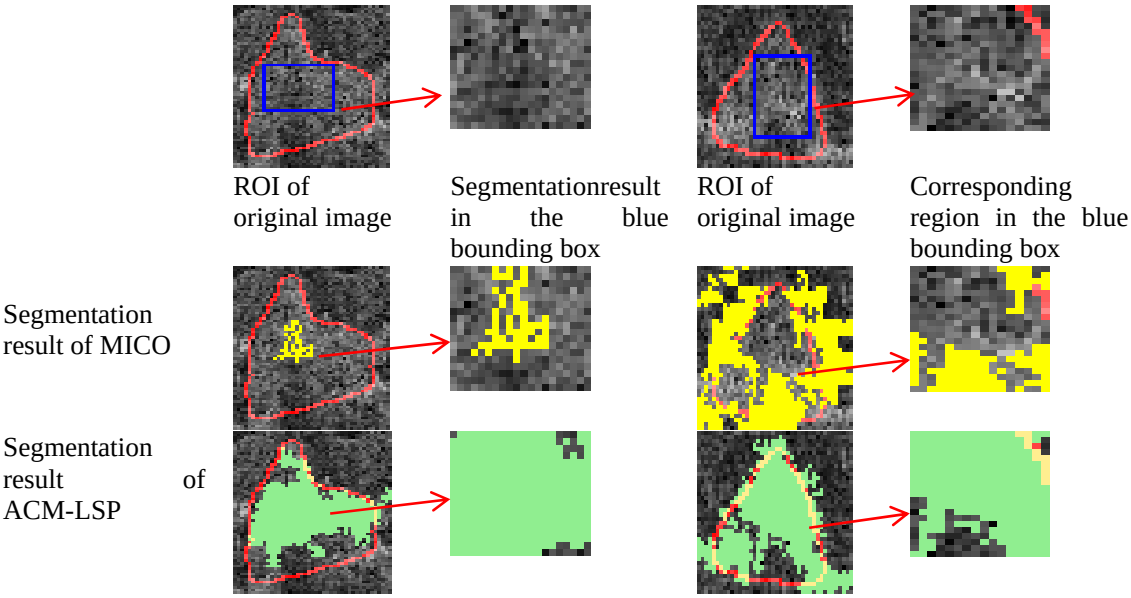


Fig.4 Segmentation result of MICO and ACM-LSP in regions with large intensity variation. The first row shows two ROI of image and the corresponding local region in the blue bounding box. The second row and the third row show the segmentation results of MICO and ACM-LSP for corresponding images respectively.

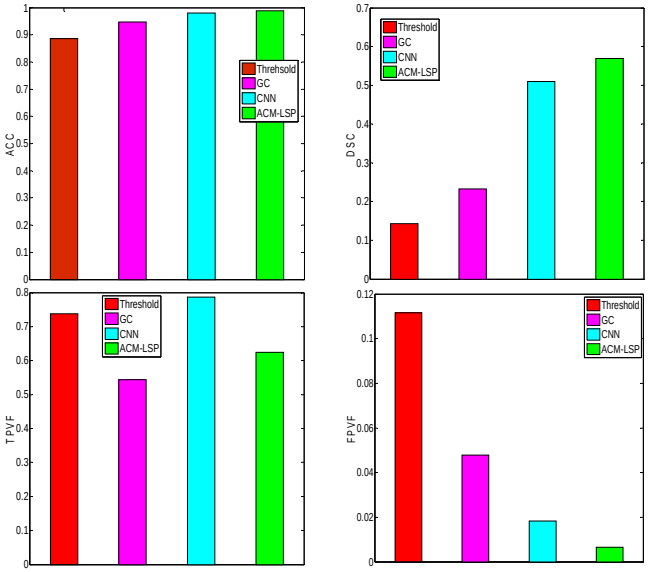


Fig.5. Segmentation method of different methods

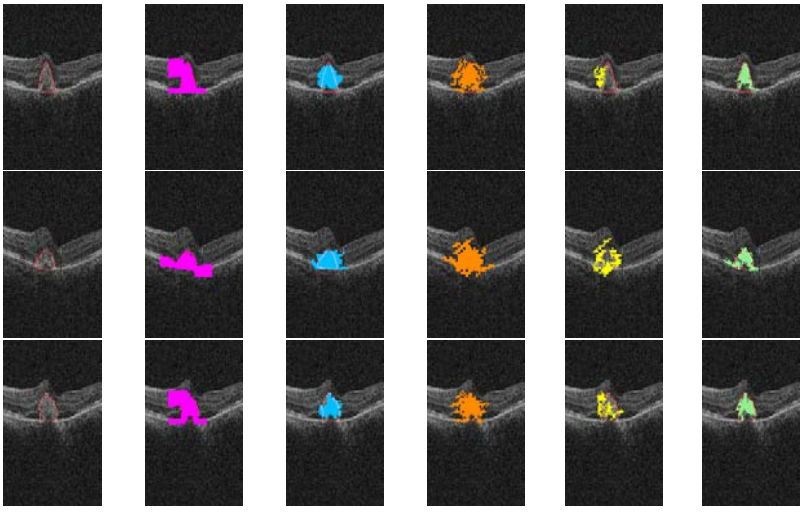
2.4 Comparison with other segmentation methods

In this experiment, we also compare ACM-LSP with state-of-the-art segmentation methods such as threshold based method, graph cut based method(GC), convolutional neural networks(CNN).

Table 3 CNV segmentation results, Compared with other method

	ACC	DSC	TPVF	FPVF
Threshold	0.8861+0.0218	0.1426+0.0238	0.7372+0.0825	0.1116+0.0201
GC	0.9471+0.0331	0.2319+0.0679	0.5435+0.1407	0.0477+0.0322
CNN	0.9797+0.0051	0.5092±0.074	**0.7862±0.045**	0.0183±0.005
ACM-LSP	**0.9878+0.004**	**0.5692+0.0484**	0.6239+0.1381	**0.0065+0.002**

Table 3 and Fig.5show ACC,DSC, TPVF, FPVF of different methods. Fig.6 gives several segmentation examples of these methods. As observed in these figures, CNN outperforms than threshold based method and GC, obtaining the best TPVF. However, ACM-LSP achieves the best ACC, DSC, FPVF, especially for DSC, increasing about 6 percentage points compared with CNN.



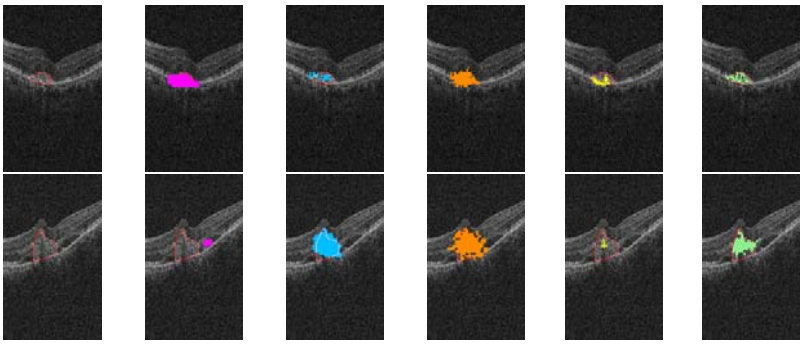


Fig.6 segmentation examples of different methods

For traditional segmentation methods, it's difficult for Threshold based method and GC to deal with complicated intensity distribution effectively. CNN has achieved promising performance in some tasks due to its powerful learning ability. However, its performance is not satisfied in our task due to the confused distribution of training patches. Fig.7 gives the similarity matrix of patches. In this figure, C_i denotes the CNV class from the *i*-th OCT data while B_i denotes the background class from the *i*-th OCT data. For each class, a center is used to represent the class and is calculated as the average of the patches in this class. And the similarity matrix is obtained by calculating the similarity between two arbitrary centers. Euclidean distance is used as the similarity measure. In this matrix, small values indicate that two corresponding classes are similar while large values indicate that two corresponding classes are different. As shown in this figure, we can infer that large intra-class variation and inter-class similarity exist in the extracted patches. For example, the similarity between C_1 and B_i is larger than the similarity between C_1 and C_2. Therefore, based on the confused distribution of training instances, it's difficult to learn an effective CNN model in our task, resulting in performance degradation. The proposed method introduces learned local similarity prior by considering the similarity relationship between neighboring pixels, which is more robust to the intra-class intensity variations and inter-class intensity similarities. Therefore, ACM-LSP achieves better performance than traditional methods.

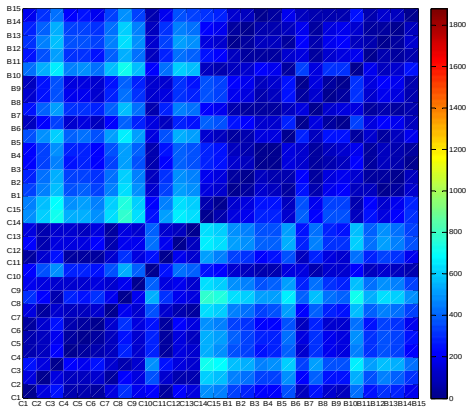


Fig.7. Similarity matrix of patches extracted from OCT images in our database