



Fast esophageal layer segmentation in OCT images of guinea pigs based on sparse Bayesian classification and graph search

CONG WANG,^{1,2,3} MENG GAN,^{2,3} NA YANG,¹ TING YANG,¹ MIAO ZHANG,¹ SIHAN NAO,¹ JING ZHU,¹ HONGYU GE,¹ AND LIRONG WANG^{1,*}

¹*School of Electronic and Information Engineering, Soochow University, Suzhou 215006, China*

²*Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China*

³*These authors contributed equally to this work and should be considered co-first authors*

**sudaweipa@163.com*

Abstract: Endoscopic optical coherence tomography (OCT) devices are capable of generating high-resolution images of esophageal structures at high speed. To make the obtained data easy to interpret and reveal the clinical significance, an automatic segmentation algorithm is needed. This work proposes a fast algorithm combining sparse Bayesian learning and graph search (termed as SBGS) to automatically identify six layer boundaries on esophageal OCT images. The SBGS first extracts features, including multi-scale gradients, averages and Gabor wavelet coefficients, to train the sparse Bayesian classifier, which is used to generate probability maps indicating boundary positions. Given these probability maps, the graph search method is employed to create the final continuous smooth boundaries. The segmentation performance of the proposed SBGS algorithm was verified by esophageal OCT images from healthy guinea pigs and the eosinophilic esophagitis (EoE) models. Experiments confirmed that the SBGS method is able to implement robust esophageal segmentation for all the tested cases. In addition, benefiting from the sparse model of SBGS, the segmentation efficiency is significantly improved compared to other widely used techniques.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Optical coherence tomography (OCT) as a powerful medical imaging technique can generate high-resolution cross-section images of biological tissues non-invasively [1]. Initial applications of OCT were mainly in ophthalmology and have achieved numerous success in retinal disease diagnosis [2–4]. By combining fiber-optic flexible endoscopes, OCT is able to enter the upper gastrointestinal tract [5, 6]. In that case, it can image the microscopic structure of the esophagus, which helps to diagnose a variety of esophageal diseases, such as Barrett's esophagus (BE) [7, 8], eosinophilic esophagitis (EoE) [9], and dysplasia [10]. Esophageal diseases are identified from OCT images by observing abnormalities in tissue microstructures, such as changes in layer thickness or disruption to the layers. For instance, BE has an irregular mucosal surface and may present an absence of the layered architecture [7]; EoE is often featured with increased basal zone thickness [9]. Finding these diseased characteristics by observing or manual segmentation is time-consuming and requires specific domain knowledge. As a result, a computer-aided layer segmentation method is urgently needed, as employed in ophthalmology [11, 12].

Representative segmentation techniques for OCT images can be grouped into two main categories. The first group is based on mathematical models. Methods of this kind build a mathematical model based on apriori assumptions of the input image structure and locate boundaries by optimizing the model. Typical methods in this class include the A-scan based

methods [2, 3, 13], the active contour [4, 14–17] and graph based methods [18–22]. The second category is machine learning based methods. Such methods formulate the segmentation into a classification problem. Features reflecting layer structure characteristics are extracted for the training, and boundaries of new images are located by pixel classifying using the trained model. Following this process, Vermeer et al. used the Haar-like features and the Support Vector Machine (SVM) to implement retinal segmentation [23]. Lang et al. employed the random forest (RF) classifier and 27 features that provide spatial, local and context awareness to segment eight retinal layers [24]. Recently, the newly developed deep learning technique, especially convolutional neural networks attracts wide attentions in the field of medical image analysis and has achieved great success in OCT image segmentation [25–28]. Different from traditional machine learning techniques, deep learning is able to learn optimal features directly from the training data and achieves higher classification accuracy [29]. However, deep learning based segmentation relies on a large number of annotated data and suffers from high computational burdens [25].

Most existing reports on esophageal OCT image segmentation employed the model-based approach. For example, Ughi used an A-scan based method to locate the esophageal wall position [30]. Zhang et al. used the graph theory dynamic program (GTDP) to implement five-layer segmentation [31]. On this basis, Gan et al. combined modified Canny edge detection with GTDP to realize more accurate segmentation [32]. To the best of our knowledge, the GTDP is the first effective framework to segment five esophageal layers [31]. However, some problems limit its application. Firstly, GTDP requires tissues to have similar widths in the same layer across the image, otherwise it will generate a fairly biased output. In other words, GTDP is difficult to deal with the possible variations in esophageal tissues. Secondly, it takes about 6 to 10 seconds to segment a single B-scan [31, 32], which is less efficient.

To solve these problems, this study proposes a sparse Bayesian graph search (SBGS) method that combines sparse Bayesian classification [33, 34] and the GTDP for esophageal layer segmentation. Sparse Bayesian classification is a Bayesian extension of SVM [35] that trains classifiers under a series of probability assumptions. It uses a prediction model similar to SVM, but with much less "support vectors" resulting from a sparsity assumption, thus ensuring high classification efficiency. Sparsity assumption is widely adopted in signal and image processing and has been confirmed to be effective in OCT image analysis as discussed in the literature [36, 37]. Another advantage of the sparse Bayesian classifier is that the probability output can be further optimized to generate the final boundary. The proposed SBGS algorithm consists of the following steps. Firstly, representative features were extracted to delineate esophageal layer structures. In detail, this study utilizes the multi-scale gradients and averages [38] to describe intensity changes along A-scan at different scales. Besides, the Gabor wavelet features [39] are also employed as a supplementary description of boundary information in other directions. Given these features, a sparse Bayesian classifier is trained to estimate the position of esophageal layer boundaries and generate a probability map. Finally, a graph is constructed based on the probability map, which is further optimized by the GTDP to get continuous smooth boundaries that segment different tissue layers. Benefitting from sparse Bayesian classification and the GTDP, the proposed SBGS method is able to achieve fast esophageal layer segmentation and will be robust in segmenting OCT images with irregular tissue layers.

In summary, our main contribution is implementing a fast segmentation of esophageal OCT images obtained from guinea pigs. More detailedly, this work contributes to extract features that comprehensively describe esophageal tissue layers based on wavelet-related theories. Besides, an efficient boundary identification scheme was designed using sparse Bayesian representation. Finally, segmentation experiments on esophageal OCT images from guinea pigs were performed to evaluate the generalization ability and clinical potential of proposed SBGS algorithm.

The rest of this study is organized as follows. Section 2 describes the details of the proposed SBGS algorithm for esophageal layer segmentation. Section 3 presents experimental results on

esophageal OCT images from guinea pigs. Discussions and conclusions are given in Sections 4 and 5, respectively.

2. Methods

The performance of our method is evaluated by segmenting endoscopic OCT images of guinea pig esophagus. A typical OCT image is shown in Fig. 1. The tissue layers marked in the images are the epithelium stratum corneum (SC), epithelium (EP), lamina propria (LP), muscularis mucosae (MM) and submucosa (SM), which are defined by boundaries B_1 to B_6 . Our SBGS

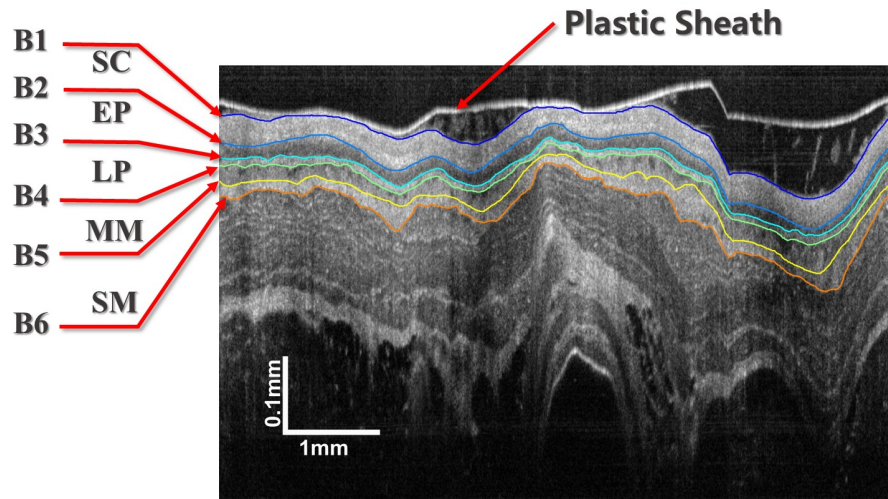


Fig. 1. A manual segmented esophageal OCT image from the guinea pig.

algorithm intends to locate B_1 to B_6 fast and robustly, thus accomplishing proper segmentation of different tissue layers. The entire process is summarized in Fig. 2. Note that the algorithm is composed of four units. The preprocessing unit regularizes the data to make it appropriate for boundary identification. The second unit extracts features reflecting tissue structures. Given these features, a sparse Bayesian classifier is achieved by the classifier training unit, which can generate a probability map for the B-scan. Finally, the probability map is processed by the boundary refinement unit to obtain smooth continuous boundaries. Details implementation about these units are presented as follows.

2.1. Preprocessing

Several preprocessing steps were taken to make the data appropriate for classifier training, including denoising, plastic sheath removal and flattening. Image denoising is achieved with a simple median filter of size 7×7 considering its high efficiency and easy implementation. The effectiveness of it in OCT images denoising has been confirmed by several researches [3, 25, 40]. The sheath removal and flattening process is specifically designed for esophageal OCT images and is described below.

2.1.1. Plastic sheath removal

The plastic sheath is used to protect the probe from biofluid during the imaging process and is recorded on OCT images as shown in Fig. 1. Its boundary is so prominent that it causes strong disturbance for the detection of the esophageal lumen boundary. Therefore, it is necessary to remove such sheath and this study achieved this using the GTDP algorithm [20].

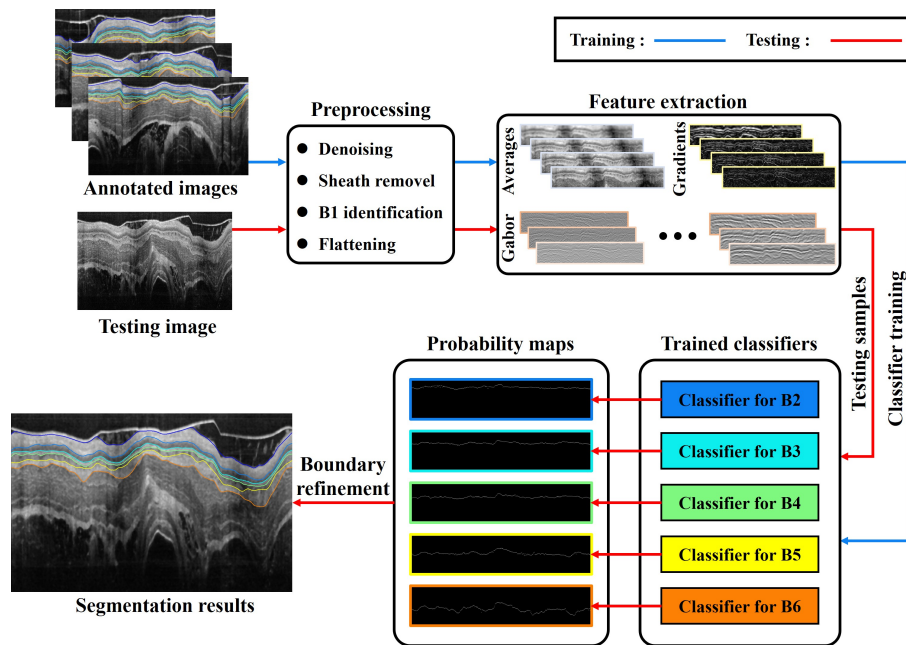


Fig. 2. Flowchart of the SBGS algorithm.

As shown in Fig. 3(a), the plastic sheath boundary P_1 often possesses the highest intensity contrast, which indicates P_1 owns the highest gradient that can be easily located by GTDP. On the basis of P_1 , P_2 can be detected by limiting the GTDP region to ten pixels below P_1 , which is an approximate sheath thickness. In that case, the plastic sheath can be removed by shifting the pixels from P_1 to P_2 and the empty pixels are filled with a mirror image. A typical result is shown in Fig. 3(b). Although there are some white regions between the plastic sheath and lumen boundary, these regions are not so prominent as the plastic sheath. As a result, the lumen boundary owns the highest gradient in this case, thus making it easily detected by the GTDP.

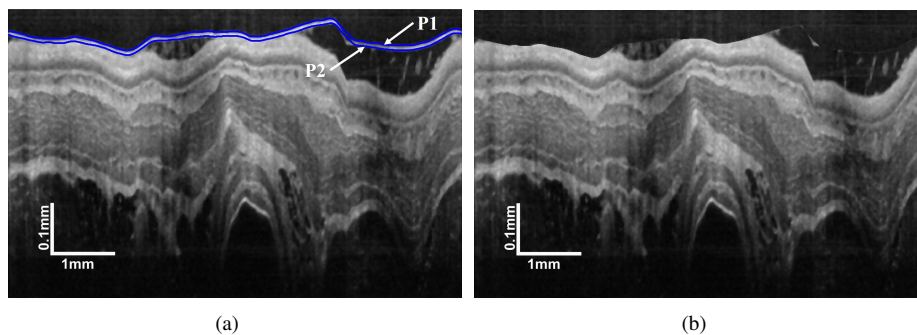


Fig. 3. Plastic sheath removal: (a) position of P_1 and P_2 ; (b) image with the plastic sheath removed.

2.1.2. Lumen boundary detection and flattening

Subsequent preprocessing is to detect the esophageal lumen boundary (B_1 in Fig. 1) and flatten the image based on it. This serves to provide spatial alignment of OCT images, which can help constrain the search area as well as reducing the algorithm sensitivity to tissue curvature.

Lumen boundary detection is convenient using the GTDP because it is the most prominent layer boundary on an image without the plastic sheath (3(b)). As a result, it can be detected by the same process as P_1 identification. The lumen boundary is then flattened by translating the A-scans up and down. In that case, OCT images are regularized for the following segmentation as shown in Fig. 4. To reduce the memory size and improve the algorithm efficiency, the layer

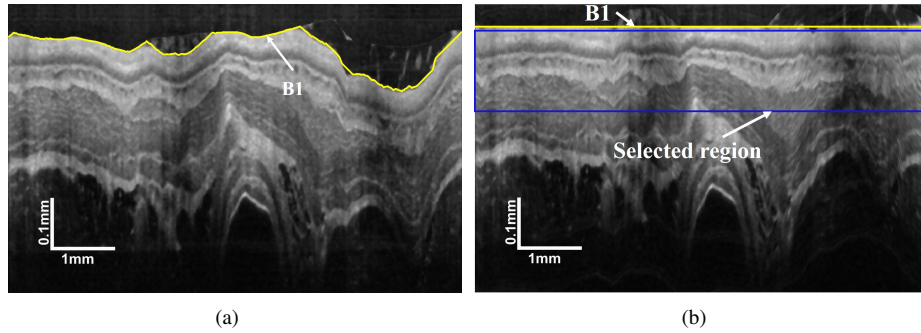


Fig. 4. Demonstration of (a) position of B_1 and (b) the flattened OCT image with marked selected region.

search area is constrained to the tissue-related regions. The total thickness of the five tissue layers for the guinea pig is $110 \mu\text{m}$ (125 pixels) in healthy condition and $150 \mu\text{m}$ (170 pixels) for the EoE model [31, 32]. Considering irregularities, limiting the search region from 5 pixels to 195 pixels below the lumen boundary (Fig. 4(b)) is sufficient for analyzing all those tissue layers.

2.2. Feature extraction

An interface occurs as the OCT signal increases or decreases, resulting in intensity changes of pixels above and below the boundary. On this basis, two kinds of features were used in this study. Features of the first type are the multi-scale gradients and averages [38], which can reflect boundary information at different resolutions. Features of the second type are the Gabor wavelet coefficients, which provides boundary information in different orientations.

2.2.1. Multi-scale gradients and averages

Since the esophageal images have been flattened, gradients and averages along A-scan are more useful for layer boundary detection. Let $I(y)$ denote the intensity along A-scan, the averages h^d at scale d are defined by

$$h^d(z) = \frac{1}{2^d} \sum_{\Delta y=1-2^{d-1}}^{1+2^{d-1}} I(y + \Delta y) \quad (1)$$

Accordingly, the gradients g^d at scale d can be calculated by

$$g^d(y) = h^d(y + 2^{d-1}) - h^d(y - 2^{d-1}) \quad (2)$$

Specifically, $h^0(y) = I(y)$, $g^0(y) = I(y + 1) - I(y)$. Multi-scale gradient and average features for pixel y can be summarized as

$$f_{ga} = [h^0(y), g^0(y), h^1(y), g^1(y), \dots, h^d(y), g^d(y)] \quad (3)$$

In this study, d is set at 4, indicating 8 features will be achieved. Typical gradient and average maps in different scales are shown in Fig. 5. It can be found that a lower scale feature map provides more precise description of local edges, whereas a higher scale feature map is more robust to noise and local distortion.

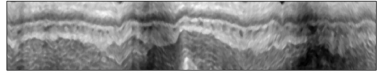
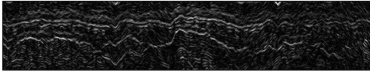
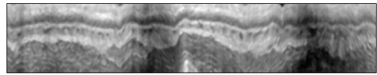
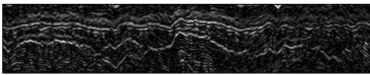

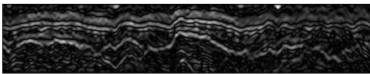

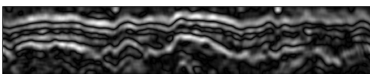
Scale	Averages	Gradients
0		
1		
2		
3		

Fig. 5. Demonstration of intensity averages and gradients at different scales.

2.2.2. Gabor wavelet features

Multiscale gradients only describe intensity changes along the A-scan. To provide information for layer boundaries in different directions, the Gabor wavelet features with directional selectivity were employed in this study.

The 2-D wavelet transform of an image I is defined as:

$$WI(a, \theta, x_0, y_0) = \|a\|^{-1} \int \int I(x, y) \psi_{\theta} \left(\frac{x - x_0}{a}, \frac{y - y_0}{a} \right) dx dy \quad (4)$$

where ψ_{θ} is the 2-D mother wavelet, a denotes the scaling parameters, x_0 and y_0 represents the spatial shifting, and θ is the orientation parameter. The Gabor function can be used as the mother wavelet, which is defined by Eq. (5) [39],

$$G(x, y) = \frac{f^2}{\pi \gamma \eta} \exp \left(-\frac{x_r^2 + \gamma^2 y_r^2}{2\sigma^2} \right) \exp(j2\pi f x_r + \phi) \quad (5)$$

$$x_r = x \cos \theta + y \sin \theta$$

$$y_r = -x \sin \theta + y \cos \theta$$

where f is the frequency of the sinusoid, θ is the orientation of the major axis of the Gaussian, ϕ is the phase offset, σ is the standard deviation of the Gaussian and γ is a ratio parameter. The detailed implementation of the Gabor wavelets can be found in [39].

In this study, a family of $U \times V$ Gabor wavelets is required to perform the multi-resolution analysis, which is defined as

$$\{G_{f_u, \theta_v}(x, y)\}$$

$$f_u = \frac{f_{\max}}{\sqrt{2}^u}, u = 0, 1, \dots, U - 1 \quad (6)$$

$$\theta_v = \frac{v}{V} \pi, v = 0, \dots, V - 1$$

where f_u and θ_v defines the orientation and scale of the Gabor wavelets, f_{\max} is the maximum central frequency.

An example of Gabor wavelet coefficients at 4 scales and 8 orientations can be found in Fig. 6 with $f_{\max} = \pi/2$, $\sigma = \pi/3$, $\gamma = 1$. Considering the fact that the data were previously flattened, the wavelets with $v = \{0, 1, 7\}$ are sufficient for describing the layer boundaries as marked in Fig. 6. As a result, 12 Gabor features for each pixel will be obtained. Demonstration of Gabor wavelet features at different scales and orientations can be found in Fig. 7.

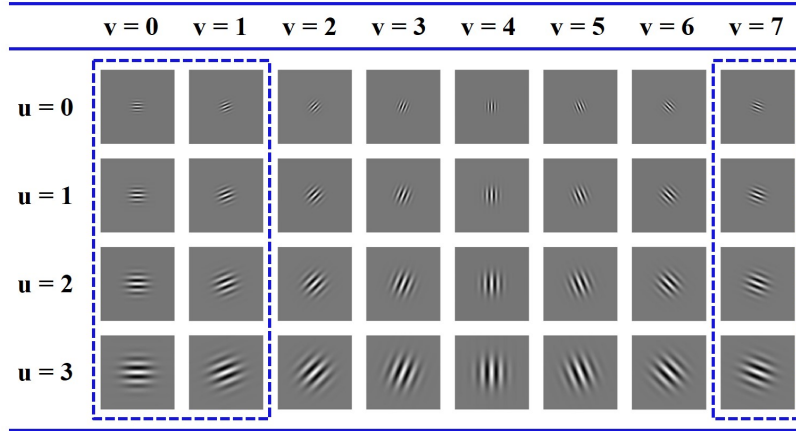


Fig. 6. Demonstration of the Gabor wavelet coefficients.

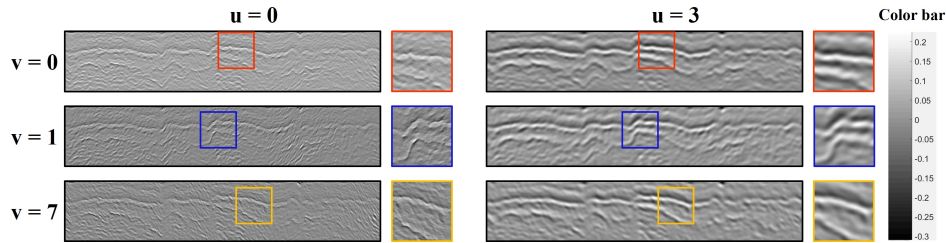


Fig. 7. Demonstration of Gabor wavelet features for esophageal OCT images.

Following the process above, we achieved 8 multi-scale gradient and average features and 12 Gabor wavelet features, totally 20 features for the subsequent boundary classification.

2.3. Sparse Bayesian classification

Suppose the data is denoted by $\{\mathbf{x}_n, t_n\}_{n=1}^N$, where \mathbf{x}_n is the feature vector, $t_n \in \{0, 1\}$ is the label for binary classification. Sparse Bayesian learning models the data based on the Bernoulli distribution, which indicates the likelihood can be expressed as

$$P(\mathbf{t}|\mathbf{w}) = \prod_{n=1}^N \sigma\{y(\mathbf{x}_n; \mathbf{w})\}^{t_n} [1 - \sigma\{y(\mathbf{x}_n; \mathbf{w})\}]^{1-t_n} \quad (7)$$

where $\sigma(x) = 1/(1 + e^{-x})$ is the sigmoid function, $y(\mathbf{x}; \mathbf{w})$ can be calculated by

$$y(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^N w_i K(\mathbf{x}, \mathbf{x}_i) + w_0 \quad (8)$$

$K(\cdot, \cdot)$ is the kernel function, and the Gaussian Kernel is employed in this study as shown in Eq. (9).

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp \{-\gamma \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2\} \tag{9}$$

To make the weight vector sparse, a zero-mean Gaussian prior distribution is set over \mathbf{w} :

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^N \mathcal{N}(w_i|0, \alpha_i^{-1}) \tag{10}$$

Following the assumptions described by Eqs. (7) to (10), an optimal $\hat{\mathbf{w}}$ can be obtained using the Laplace's method [41]. For a new sample $\tilde{\mathbf{x}}$, the probability for $\tilde{t} = 1$ can be calculated by

$$p(\tilde{t} = 1|\mathbf{w}) = \sigma\{y(\tilde{\mathbf{x}}; \mathbf{w})\} = \sigma\left[\sum_{i=1}^{N_r} w_i K(\mathbf{x}, \mathbf{x}_i) + w_0\right] \tag{11}$$

where \mathbf{x}_i indicates the relevance vector, which has nonzero w_i , N_r is the relevance vector number. Since \mathbf{w} is sparse, only a few \mathbf{x}_i are needed for a new sample identification, thus making the sparse Bayesian classification quite efficient.

The sparse Bayesian classifier using the Bernoulli model is a binary classifier, which can be generalized to deal with multi-class problems by the "one-versus-all" strategy [42] or "one-versus-one" strategy [42]. The main purpose of this study is to decide if a pixel belongs to any of the five boundaries. In this case, the "one-versus-all" strategy is more appropriate, separating the pixels on one boundary from the rest. Using this strategy, we trained five binary sparse Bayesian classifiers to detect pixels on boundaries B_2 to B_6 , respectively as shown in Fig. 2. To achieve the classifier for B_i , points on the i -th boundary are treated as the positive samples (labeled "1"), and the remainders in the training set are regarded as the negative samples (labeled "0"). In this way, we can model the data by Eq. (7) and achieve a binary sparse Bayesian classifier for boundary B_i .

It is worth mentioning that the sparse Bayesian classifier can achieve multi-classification by modeling the data with multinomial distribution [33], but it suffers from a highly disadvantageous consequence from a computational perspective [33].

2.4. Boundary identification based on the classification result

We marked the pixel with the highest probability in each A-scan to get the most likely boundary position. The boundary identification process can be depicted by Fig. 8. Note that each pixel was represented by a feature vector, which is then individually classified to achieve the final probability map. A typical result was shown in Fig. 9. Although the boundaries can be visually

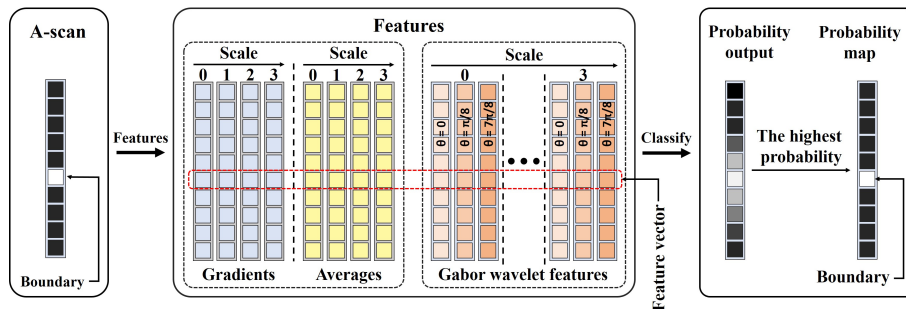


Fig. 8. Flowchart of the boundary identification process.

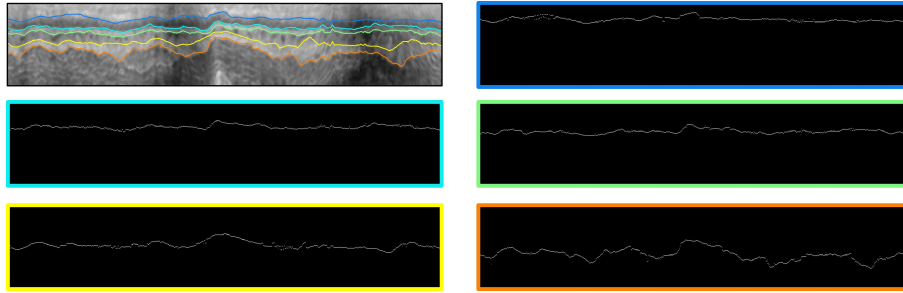


Fig. 9. Presentation of the boundary position indicated by sparse Bayesian classification. The probability map has a border painted with the same color as the corresponding boundary shown by the image on top left.

identified, there are still some missing parts or dropouts. As a result, the probability map needs to be further refined.

An intuitive way is to smooth the boundaries shown in Fig. 9 and inpaint the missing points. This method is simple and fast, but requires the classifier to be accurate and is susceptible to noise or other disturbance. A more robust approach is using the GTDP as in the references [24, 25]. This approach constructs a graph based on the probability map rather than the gradients and finds the optimal path by Dijkstra's shortest path algorithm [20]. To refine the probability map for boundary B_i , $i = 2, \dots, 6$, the weight for edge connecting adjacent pixels a and b is defined as

$$W_{ab}^i = 2 - (P_a^i + P_b^i) + w_{\min} \quad (12)$$

where P_a^i and P_b^i are probability map intensities of B_i for points a and b . w_{\min} is the minimum possible weight in the graph. More details about GTDP can be found in Chiu's work [20]. It is worth mentioning that the search region for each boundary is ± 6 pixels around the most probable boundary position, which will make the optimization more efficient. The region width is set based on the fact that the smallest tissue layer has a thickness of around $10 \mu\text{m}$ (11 pixels) [31, 32]. In that case, searching a certain boundary in ± 6 pixels around the most probable position utilizes as much information as possible without inducing inference from boundaries of other tissue layers.

3. Experiments

3.1. Data and experimental environment

Esophageal OCT images from guinea pigs were obtained using an 800-nm ultrahigh resolution endoscopic OCT system [6, 43, 44]. Data from 4 subjects, including two healthy guinea pigs and two EoE models [45] were used to evaluate the proposed SBGS algorithm. For each subject, two datasets were obtained by the OCT system from different (deep and shallow) positions of the esophagus. Each dataset has 30 continuous imaged B-scans of size 2048×2048 . More details about the data are listed in Table 1. The data was manually annotated by an experienced grader using ITK-SNAP [46]. Manual segmentation results were used for classifier training and algorithm evaluation. In the following experiments, the B-scans were resized by 0.5, which indicates the new size of B-scans is 1024×1024 . Our images were analyzed by MATLAB on a personal computer with an Intel Core i7 2.20 GHz CPU and 16 GB RAM.

In the first part of the experiment, we performed the segmentation on each case separately, and took Case 1 as a representative example to numerically evaluate the SBGS algorithm in accuracy and efficiency. Since OCT images are collected continuously and consecutive frames have similar structure, we selected two well-imaged B-scans that roughly trisect Case 1 for

Table 1. Information of the dataset used in this study.

Datasets	Subject 1		Subject 2		Subject 3		Subject 4	
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Condition	Healthy	Healthy	Healthy	Healthy	EoE	EoE	EoE	EoE
Position	Deep	Shallow	Deep	Shallow	Deep	Shallow	Deep	Shallow
Frames	30	30	30	30	30	30	30	30

training to ensure that our system does not become biased. The remaining 28 B-scans were used for testing. To show the generation and clinical potential of the proposed method, another experiment is performed using different training and testing sets. The training sets are composed of the four cases obtained from Subjects 1 and 3, each with 2 selected B-scans. The testing sets are composed of all the frames in Subjects 2 and 4. In this case, there was no overlap between subjects used for training and testing.

3.2. Accuracy analysis of the SBGS algorithm using Case 1

As is discussed above, the training and testing sets contain two B-scans and 28 B-scans respectively in this case. For a given B-scan of size 1024×1024 , we used all the pixels on the boundaries (5 boundaries $\times 1024$ points per boundary) for training. Since the number of non-boundary pixels is much larger than the boundary pixels, we randomly select 2048 non-boundary pixels to balance the training data.

Using the trained sparse Bayesian classifier, the segmentation performance of SBGS can be evaluated by the test data. We first made an intuitive comparison between SBGS and the widely used GTDP algorithm as shown in Fig. 10. Figures 10(a) and 10(c) illustrate the SBGS segmentation results for OCT image with regular and irregular tissues, respectively. It can be found that the SBGS accomplished satisfactory segmentation in both cases. Specifically, we

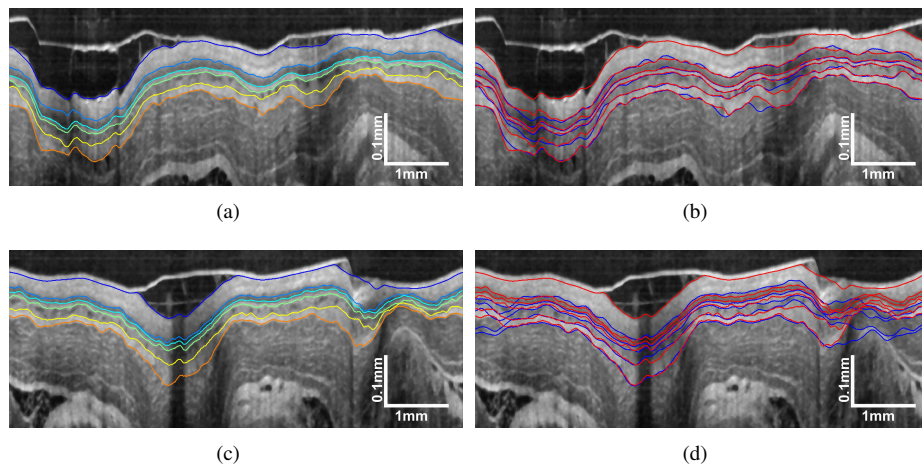


Fig. 10. Demonstration of (a) SBGS segmentation result for OCT image with regular tissues; (b) comparisons of SBGS (red line) and GTDP (blue line) results; (c) SBGS segmentation result for OCT images with tissue irregularities and (d) comparisons of SBGS (red line) and GTDP (blue line) results.

plotted the boundaries obtained by GTDP and SBGS in blue and red, respectively. In Fig. 10(b) with regular tissues, it can be visually identified that the red lines generated by SBGS match the layer structures better. In Fig. 10(d) with larger thickness variations, the blue lines identified by GTDP show a great bias from the correct tissue structure while the SBGS can still identify the correct boundaries.

To further evaluate the segmentation accuracy of the SBGS, we replaced the sparse Bayesian classifier with other classifiers widely used in OCT segmentation, including the SVM [23] and RF [24], thus achieving the SVM-GS and RF-GS methods. Parameters for different classifiers were set as follows. The sparse Bayesian learning and SVM utilized the Gaussian kernel described by Eq. (9) with $\gamma = 0.3$. For SVM, the C -Support Vector Classification was employed [47] and the regularization parameter C is set at 4000 to achieve fewer support vectors. The main parameter for RF is the number of trees, which is set at 500 in this study.

The quantitative evaluation was carried out by comparing the segmentation result of SBGS with SVM-GS, RF-GS and GTDP [20] using the manual segmentation as a reference. The unsigned border position differences between the automatic and manual segmentation for the 28 testing images from Case 1 are listed in Table 2. Results are presented in the form of mean \pm standard deviation in micrometer, with bold data indicating the result with the smallest average difference. B_1 as the baseline for flattening was identified by the GTDP in all algorithms, hence

Table 2. Unsigned border position differences of different esophageal OCT image segmentation algorithms.

Boundaries	SBGS	SVM-GS	RF-GS	GTDP
B_1 (μm)	2.46 ± 0.88	2.48 ± 0.91	2.44 ± 0.73	2.47 ± 0.93
B_2 (μm)	2.07 ± 1.66	3.72 ± 2.02	1.91 ± 1.53	4.58 ± 2.32
B_3 (μm)	0.22 ± 0.80	1.21 ± 1.33	0.31 ± 1.09	1.92 ± 1.89
B_4 (μm)	0.66 ± 0.97	1.87 ± 1.67	1.10 ± 1.12	2.11 ± 2.08
B_5 (μm)	3.44 ± 2.23	4.11 ± 3.28	3.22 ± 2.34	5.03 ± 4.06
B_6 (μm)	1.31 ± 1.08	2.20 ± 1.84	2.05 ± 1.69	2.51 ± 2.10

the error values are almost identical. Except for B_1 , classifier-based methods (SBGS, SVM-GS, RF-GS) achieved more accurate boundary identification than the GTDP since they synthesized more comprehensive features to delineate the tissue structure. Besides, SBGS shows higher accuracy than SVM-GS because it uses fewer vectors to prevent overfitting and does not require optimization of the additional parameter C [47]. The segmentation result of RF-GS is comparable to SBGS in accuracy. In detail, SBGS achieved the most accurate segmentation in B_3 , B_4 and B_6 , while the RF-GS has higher accuracy in segmenting B_2 and B_5 .

3.3. Efficiency analysis of the SBGS algorithm using Case 1

To evaluate the efficiency of the proposed algorithm, we listed the average segmentation time of SBGS, SVM-GS, RF-GS and GTDP in Table 3 for the testing data from Case 1. It is noted that the SBGS method is significantly faster than the other classifier-based methods. The main reason is that the sparse Bayesian classifier has a linear prediction model [Eq. (8)] that depends on much fewer relevance vectors as shown in Table 4. As a result, the classification can be implemented by a few linear operations, which is quite efficient. Moreover, the graph search is performed within a small region as illustrated in Section 2, thus making SBGS is also faster than GTDP.

For an intuitive explanation of why a sparse model is beneficial for efficient segmentation, we analyzed the detailed classification process of the sparse Bayesian classifier and the SVM. The classification of a new input for SVM depends on the support vectors (SVs), and the sparse

Table 3. Computation time for OCT image segmentation.

Methods	SBGS	SVM-GS	RF-GS	GS
Time (s)	1.71 ± 0.13	12.24 ± 2.08	7.82 ± 1.11	10.01 ± 1.84

Bayesian classifier has similar vectors, which are called relevance vectors (RVs) as shown in Eq. (11). Accordingly, the prediction efficiency is determined by the vector numbers for these two classifiers. In our segmentation framework, a B-scan generates $1024 \times 5 + 2048 = 7168$ pixels for training, each with 20 features. The classifier is trained using 2 representative B-scans, resulting in a training set of size 14336×20 . The detailed analysis about the SVM and sparse Bayesian classifiers can be found in Table 4, where N_r represents the number of SVs and RVs used for classification. The sparse Bayesian classification achieved much smaller N_r than SVM, which indicates the boundary identification can be accomplished based on only a few tens of samples. As a result, the computation time can be remarkably reduced from around 1.5 to 0.15 seconds for recognizing a single boundary.

Table 4. Comparisons of N_r and the classification time for SVM and sparse Bayesian classifier (denoted by SBC).

Boundaries	N_r		Times (s)	
	SVM	SBC	SVM	SBC
B2	834	30	1.77 ± 0.07	0.12 ± 0.01
B3	612	32	1.22 ± 0.06	0.13 ± 0.01
B4	677	38	1.47 ± 0.05	0.14 ± 0.02
B5	1043	39	2.08 ± 0.08	0.16 ± 0.02
B6	895	31	2.12 ± 0.07	0.13 ± 0.01

3.4. Generalization ability and clinical potential

To show the generalization ability and clinical potential of the SBGS algorithm, the training process is performed using images from a healthy guinea pig (Subject 1) and an EoE model (Subject 3), while the testing set is composed of Subjects 2 and 4. Detailedly, we selected two B-scans from each case in Subjects 1 and 3 using the same method as described in Section 3.1 to create the training set with eight B-scans. All of the frames obtained from Subjects 2 and 4 are used for testing, making the testing set of size $30 \text{ frames/case} \times 4 \text{ cases} = 120 \text{ frames}$.

Representative segmentation results for the four testing cases can be found in Fig. 11. The tissue layers are well delineated for all the cases, which confirm that the SBGS algorithm is effective even if the layer structure changes due to such pathology. In addition, in this case, the training and testing sets are composed of images from different subjects, which indicates there is no overlap between training and testing, thus verifying the generalization ability of our method. Since EoE is often featured with increased basal zone thickness [9], we calculated the layer thicknesses based on the segmentation results. The corresponding statistic plot for layer thickness is presented in Fig. 12. Obviously, the SC layer thickness for EoE model is significantly larger than the normal one, thereby proving that the SBGS algorithm would potentially aid clinical diagnosis.

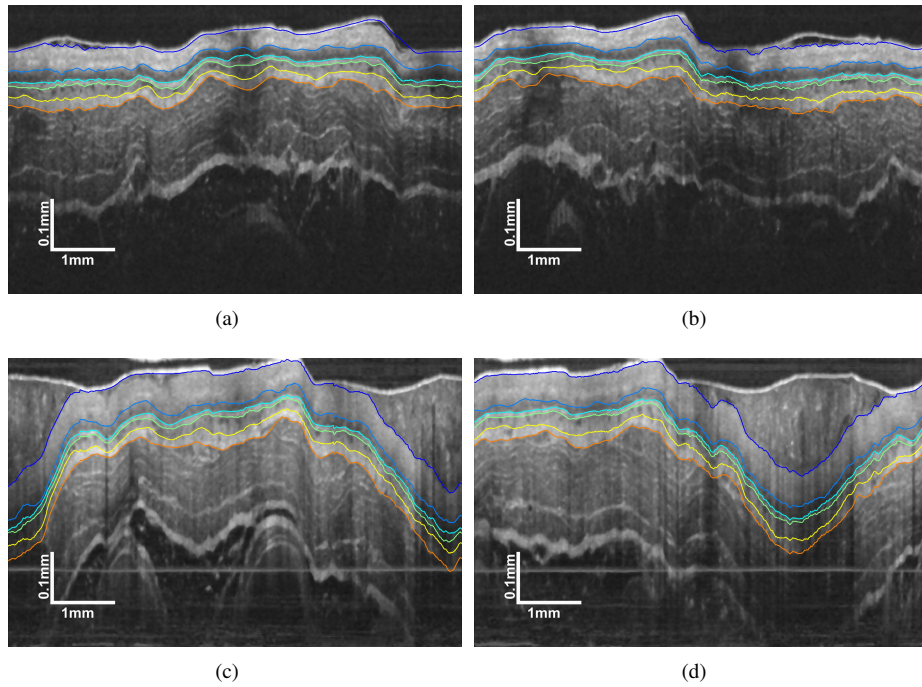


Fig. 11. Representative segmentation result by SBGS of (a) Case 3 and (b) Case 4 from the testing healthy guinea pig (Subject 2); (c) Case 7 and (d) Case 8 from the testing EoE model (Subject 4).

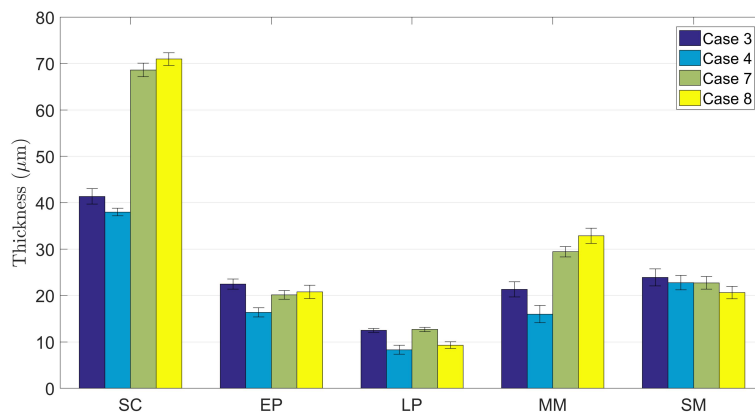


Fig. 12. Statistical results of layer thicknesses from different guinea pigs.

4. Discussions

This study used the sparse Bayesian classification for rapid esophageal layer segmentation in OCT images. The sparse Bayesian classifier offers some quite compelling and beneficial features. Firstly, the trained model is highly sparse, indicating that the prediction can be made based on a small number of samples, thus guaranteeing high classification efficiency. Besides, the noise variance can be automatically estimated. As a result, regularization parameters are no longer needed, such as C in SVM. Moreover, sparse Bayesian classification can provide a posterior probability of the sample, which is important for the following analysis, as discussed in Section 2.

The primary disadvantage of the sparse Bayesian method is the computational complexity of the training process [33]. The update rules required memory and computation scale respectively with the square and cube of the number of basis functions. In that case, sparse Bayesian learning may be less practical if the training sample number is large. Fortunately, since we only need the boundary pixels and a few background pixels for training, the model can be successfully learned. Sparse Bayesian classification needs longer training time compared to SVM and RF. In this study, sparse Bayesian classification takes about 10 minutes for model training with two B-scans, while SVM and RF only need 2 minutes and 1 minute, respectively.

Using the proposed algorithm, segmenting a single B-scan takes about 1.7 seconds with Matlab. The algorithm can be sped up by further optimizations. For example, a faster programming language, such as C can be employed. Besides, the classification of each pixel can be parallelly processed by multiple computer cores or the graphics processing unit (GPU). In that case, a more efficient segmentation system can be built, which can meet the requirements of online segmentation.

Experiments have proved that the SBGS algorithm is able to delineate EoE OCT data, suggesting that it has the potential to be extended for other esophageal diseases. Taking BE as an example, the OCT image of BE may present an absence of layered architecture. In that case, the SBGS is expected to implement tissue layer segmentation by extracting different features or adding a new criterion to locate regions with missing layers. Benefiting from the learning ability of sparse Bayesian classification, the SBGS will detect the pathological changes of the OCT image based on the annotated data, thus achieving automatic disease diagnosis.

It is worth mentioning that the deep learning based methods are very popular in image analysis, and has been introduced to retinal OCT image processing [25–28] recently. However, for esophageal OCT images discussed in this study, the deep network is hard to train due to the limited annotated samples. To verify this, we performed an additional experiment based on a representative deep learning based segmentation method proposed by Fang et al. [25] using the same training and testing sets as Section 3.2. The main idea is using a convolutional neural network (CNN) to extract features of specific layer boundaries and classifying the six layer boundaries on esophageal OCT image. According to Fang's work [25], we extracted patches of size 33×33 pixels centered on each pixel of six annotated boundaries from OCT B-scans and marked them with labels 1-6. Besides, one randomly selected patch centered at the background for each A-scan is labeled with 0. In this case, the two B-scans in training set generated 14336 patches. The CNN architecture used in this experiment was summarized in Table 5. It can be calculated that the number of parameters for this network is 1,625,479, which is difficult to optimize using the 14,336 patches. In this case, after 12 epochs optimization with batch size 128, the training accuracy is 0.7726 that cannot generate useful probability maps for boundary identification. When we add the training B-scans from 2 (14,336 patches) to 12 (86,016 patches), the training accuracy reached 0.9446 with the same configuration. It can be found that the CNN needs more annotated samples to optimize the huge number of parameters, which is in coincident with the discussion in [25]. In contrast, the SBGS algorithm showed effectiveness even if the train image is limited as discussed in Section 3.2.

Table 5. The CNN architecture used in this experiment.

Layers	Type	Filter Size	Neuron Numbers	Output Shape
Layer 1	Convolution	3×3	32	$31 \times 31 \times 32$
Layer 2	Convolution	3×3	64	$29 \times 29 \times 64$
Layer 3	Max pool	–	–	$14 \times 14 \times 64$
Layer 4	Dropout	–	–	$14 \times 14 \times 64$
Layer 5	Flatten	–	–	12544×1
Layer 6	Dense	–	128	128×1
Layer 7	Dropout	–	–	128×1
Layer 8	Dense	–	7	7×1

5. Conclusions

In this study, we presented an SBGS method that combines sparse Bayesian learning and graph search for fast segmentation of layer boundaries on esophageal OCT images. The SBGS algorithm utilizes sparse Bayesian learning to build a classifier based on multi-scale gradients, averages and Gabor wavelet coefficients and generates a probability map of boundary positions. A graph is then constructed using the probability map, which is subsequently used to generate a continuous and smooth boundary by the graph search technique. In that case, esophageal layers are clearly segmented, making the OCT image easy to interpret. Experiments on segmenting OCT images from guinea pig esophagus demonstrated that the proposed SBGS method is robust and much faster than SVM-GS, RF-GS and the GTDP. In addition, segmentation experiments on esophageal OCT images from normal guinea pigs and EoE models confirmed that the SBGS algorithm has the potential to diagnose esophageal diseases. It is also appealing to apply the SBGS to other layered structure endoscopic OCT images (like airway) and non-endoscopic OCT images (like retina) in future work.

Funding

National Key R&D Program for Major International Joint Research of China (2016YFE0107700). Jiangsu Planned Projects for Postdoctoral Research Funds of China (2018K007A, 2018K044C).

Acknowledgments

We would like to acknowledge Prof. Xingde Li and Dr. Wu Yuan from the Johns Hopkins University for their technical support in the SD-OCT systems and animal experiments.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science* **254**, 1178–1181 (1991).
2. M. R. Hee, J. A. Izatt, E. A. Swanson, D. Huang, J. S. Schuman, C. P. Lin, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography of the human retina," *Arch. Ophthalmol.* **113**, 325–332 (1995).
3. D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements from optical coherence tomography using a markov boundary model," *IEEE Transactions on Med. Imaging* **20**, 900–916 (2001).
4. D. C. Fernandez, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Opt. Express* **13**, 10200–10216 (2005).

5. G. J. Tearney, M. E. Brezinski, B. E. Bouma, S. A. Boppart, C. Pitris, J. F. Southern, and J. G. Fujimoto, "In vivo endoscopic optical biopsy with optical coherence tomography," *Science* **276**, 2037–2039 (1997).
6. J. F. Xi, A. Q. Zhang, Z. Y. Liu, W. X. Liang, L. Y. Lin, S. Y. Yu, and X. D. Li, "Diffraction catheter for ultrahigh-resolution spectral-domain volumetric OCT imaging," *Opt. Lett.* **39**, 2016–2019 (2014).
7. J. M. Ponomarev, S. Brand, B. E. Bouma, G. J. Tearney, C. C. Compton, and N. S. Nishioka, "Diagnosis of specialized intestinal metaplasia by optical coherence tomography," *Gastroenterology* **120**, 7–12 (2001).
8. S. H. Yun, G. J. Tearney, B. J. Vakoc, M. Shishkov, W. Y. Oh, A. E. Desjardins, M. J. Suter, R. C. Chan, J. A. Evans, I. K. Jang, N. S. Nishioka, J. F. de Boer, and B. E. Bouma, "Comprehensive volumetric optical microscopy in vivo," *Nat. Medicine* **12**, 1429–1433 (2006).
9. Z. Y. Liu, J. F. Xi, M. Tse, A. C. Myers, X. D. Li, P. J. Pasricha, and S. Y. Yu, "Allergic inflammation-induced structural and functional changes in esophageal epithelium in a guinea pig model of eosinophilic esophagitis," *Gastroenterology* **146**, S92 (2014).
10. M. J. Suter, M. J. Gora, G. Y. Lauwers, T. Arnason, J. Sauk, K. A. Gallagher, L. Kava, K. M. Tan, A. R. Soomro, T. P. Gallagher, J. A. Gardecki, B. E. Bouma, M. Rosenberg, N. S. Nishioka, and G. J. Tearney, "Esophageal-guided biopsy with volumetric laser endomicroscopy and laser cautery marking: a pilot clinical study," *Gastrointest. Endosc.* **79**, 886–896 (2014).
11. J. P. McLean, Y. Ling, and C. P. Hendon, "Frequency-constrained robust principal component analysis: a sparse representations approach to segmentation of dynamic features in optical coherence tomography imaging," *Opt. Express* **25**, 25819–25830 (2017).
12. J. K. Zhang, B. M. Williams, S. Lawman, D. Atkinson, Z. J. Zhang, Y. C. Shen, and Y. L. Zheng, "Non-destructive analysis of flake properties in automotive paints with full-field optical coherence tomography and 3d segmentation," *Opt. Express* **25**, 18614–18628 (2017).
13. H. Ishikawa, D. M. Stein, G. Wollstein, S. Beaton, J. G. Fujimoto, and J. S. Schuman, "Macular segmentation with optical coherence tomography," *Investig. Ophthalmol. & Vis. Sci.* **46**, 2012–2017 (2005).
14. M. Mujat, R. C. Chan, B. Cense, B. H. Park, C. Joo, T. Akkin, T. C. Chen, and J. F. de Boer, "Retinal nerve fiber layer thickness map determined from optical coherence tomography images," *Opt. Express* **13**, 9480–9491 (2005).
15. A. Yazdanpanah, G. Hamarneh, B. Smith, and M. Sarunic, "Intra-retinal layer segmentation in optical coherence tomography using an active contour approach," *Med. Image Comput. Comput. Interv. - Miccai 2009, Pt II, Proc.* **5762**, 649 (2009).
16. M. A. Mayer, J. Hornegger, C. Y. Mardin, and R. P. Tornow, "Retinal nerve fiber layer segmentation on fd-OCT scans of normal subjects and glaucoma patients," *Biomed. Opt. Express* **1**, 1358–1383 (2010).
17. S. J. Niu, L. de Sisternes, Q. Chen, T. Leng, and D. L. Rubin, "Automated geographic atrophy segmentation for SD-OCT images using region-based c-v model via local similarity factor," *Biomed. Opt. Express* **7**, 581–600 (2016).
18. M. K. Garvin, M. D. Abramoff, R. Kardon, S. R. Russell, X. D. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-d graph search," *IEEE Transactions on Med. Imaging* **27**, 1495–1505 (2008).
19. Q. Yang, C. A. Reisman, Z. G. Wang, Y. Fukuma, M. Hangai, N. Yoshimura, A. Tomidokoro, M. Araie, A. S. Raza, D. C. Hood, and K. P. Chan, "Automated layer segmentation of macular OCT images using dual-scale gradient information," *Opt. Express* **18**, 21293–21307 (2010).
20. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SD-OCT images congruent with expert manual segmentation," *Opt. Express* **18**, 19413–19428 (2010).
21. K. Yu, F. Shi, E. T. Gao, W. F. Zhu, H. Y. Chen, and X. J. Chen, "Shared-hole graph search with adaptive constraints for 3d optic nerve head optical coherence tomography image segmentation," *Biomed. Opt. Express* **9**, 962–983 (2018).
22. B. T. Soetikno, L. Beckmann, X. Zhang, A. A. Fawzi, and H. F. Zhang, "Visible-light optical coherence tomography oximetry based on circumpapillary scan and graph-search segmentation," *Biomed. Opt. Express* **9**, 3640 (2018).
23. K. A. Vermeer, J. van der Schoot, H. G. Lemij, and J. F. de Boer, "Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images," *Biomed. Opt. Express* **2**, 1743–1756 (2011).
24. A. Lang, A. Carass, M. Hauser, E. S. Sotirchos, P. A. Calabresi, H. S. Ying, and J. L. Prince, "Retinal layer segmentation of macular OCT images using boundary classification," *Biomed. Opt. Express* **4**, 1133–1152 (2013).
25. L. Y. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. T. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative and patients using deep learning and graph search," *Biomed. Opt. Express* **8**, 2732–2744 (2017).
26. F. G. Venhuizen, B. van Ginneken, B. Liefers, F. van Asten, V. Schreur, S. Fauser, C. Hoyng, T. Theelen, and C. I. Sanchez, "Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography," *Biomed. Opt. Express* **9**, 1545–1569 (2018).
27. H. Z. Fu, J. Cheng, Y. W. Xu, D. W. K. Wong, J. Liu, and X. C. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Transactions on Med. Imaging* **37**, 1597–1605 (2018).
28. S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, G. Subramanian, L. Zhang, S. Perera, J. M. Mari, K. S. Chin, T. A. Tun, N. G. Strouthidis, T. Aung, A. H. Thiery, and M. J. A. Girard, "Drunet: a dilated-residual u-net deep learning network to segment optic nerve head tissues in optical coherence tomography images," *Biomed. Opt. Express* **9**, 3244–3265 (2018).
29. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc.*

- IEEE **86**, 2278–2324 (1998).
30. G. J. Ughi, M. J. Gora, A. F. Swager, A. Soomro, C. Grant, A. Tiernan, M. Rosenberg, J. S. Sauk, N. S. Nishioka, and G. J. Tearney, “Automated segmentation and characterization of esophageal wall in vivo by tethered capsule optical coherence tomography endomicroscopy,” *Biomed. Opt. Express* **7**, 409–419 (2016).
 31. J. L. Zhang, W. Yuan, W. X. Liang, S. Y. Yu, Y. M. Liang, Z. Y. Xu, Y. X. Wei, and X. D. Li, “Automatic and robust segmentation of endoscopic OCT images and optical staining,” *Biomed. Opt. Express* **8**, 2697–2708 (2017).
 32. M. Gan, C. Wang, T. Yang, N. Yang, M. Zhang, W. Yuan, X. D. Li, and L. R. Wang, “Robust layer segmentation of esophageal OCT images based on graph search using edge-enhanced weights,” *Biomed. Opt. Express* **9**, 4481–4495 (2018).
 33. M. E. Tipping, “Sparse bayesian learning and the relevance vector machine,” *J. Mach. Learn. Res.* **1**, 211–244 (2001).
 34. M. E. Tipping, “Bayesian inference: An introduction to principles and practice in machine learning,” *Adv. Lect. on Mach. Learn.* **3176**, 41–62 (2004).
 35. V. Vapnik, “The support vector method of function estimation,” in *Nonlinear Modeling*, pp. 55–85 (Springer, 1998).
 36. L. Y. Fang, S. T. Li, X. D. Kang, J. A. Izatt, and S. Farsiu, “3-d adaptive sparsity based image compression with applications to optical coherence tomography,” *IEEE Transactions on Med. Imaging* **34**, 1306–1320 (2015).
 37. L. Y. Fang, S. T. Li, D. Cunefare, and S. Farsiu, “Segmentation based sparse reconstruction of optical coherence tomography images,” *IEEE Transactions on Med. Imaging* **36**, 407–421 (2017).
 38. P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” 2001 IEEE Comput. Soc. Conf. on Comput. Vis. Pattern Recognition, Vol 1, Proc. pp. 511–518 (2001).
 39. T. S. Lee, “Image representation using 2D gabor wavelets,” *IEEE Transactions on Pattern Analysis Mach. Intell.* **18**, 959–971 (1996).
 40. V. J. Srinivasan, B. K. Monson, M. Wojtkowski, R. A. Bilonick, I. Gorczynska, R. Chen, J. S. Duker, J. S. Schuman, and J. G. Fujimoto, “Characterization of outer retinal morphology with high-speed, ultrahigh-resolution optical coherence tomography,” *Investig. Ophthalmol. & Vis. Sci.* **49**, 1571–1579 (2008).
 41. D. J. C. Mackay, “The evidence framework applied to classification networks,” *Neural Comput.* **4**, 720–736 (1992).
 42. C. Dong and L. F. Tian, “Accelerating relevance-vector-machine-based classification of hyperspectral image with parallel computing,” *Math. Probl. Eng.* **2018** 252979 (2012).
 43. W. Yuan, J. Mavadia-Shukla, J. F. Xi, W. X. Liang, X. Y. Yu, S. Y. Yu, and X. D. Li, “Optimal operational conditions for supercontinuum-based ultrahigh-resolution endoscopic OCT imaging,” *Opt. Lett.* **41**, 250–253 (2016).
 44. W. Yuan, R. Brown, W. Mitzner, L. Yarmus, and X. D. Li, “Super-achromatic monolithic microprobe for ultrahigh-resolution endoscopic optical coherence tomography at 800 nm,” *Nat. Commun.* **8**, 1531 (2017).
 45. Z. Y. Liu, Y. T. Hu, X. Y. Yu, J. F. Xi, X. M. Fan, C. M. Tse, A. C. Myers, P. J. Pasricha, X. D. Li, and S. Y. Yu, “Allergen challenge sensitizes trpa1 in vagal sensory neurons and afferent c-fiber subtypes in guinea pig esophagus,” *Am. J. Physiol. Liver Physiol.* **308**, G482–G488 (2015).
 46. P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,” *Neuroimage* **31**, 1116–1128 (2006).
 47. C. C. Chang and C. J. Lin, “Libsvm: A library for support vector machines,” *ACM Transactions on Intell. Syst. Technol.* **2**, 27 (2011).