Original Article

Rapid training of layer specific edges for segmentation of retinal layers in OCT

S P K Karri[1], Debjani Chakraborty[2], Jyotirmoy Chatterjee[1]

Abstaract

Aim: Optical Coherence Tomography scans based retinal layer width profiles can be attributed to various pathologies and graph based methods have illustrated accurate retinal layers segmentation particularly in presence of noise.. Graph based segmentation approaches employ the edge information to infer the class at each image location (pixel). **Material and method:** Total 10 patients presented with diabetic macular edema (DME) were selected for the study. Machine learning is used for predictions of edges given an image to avoid heuristics based edge detection (image processing). Current approach aims at training one model with the capability of predicting multiple classes along with edge. **Result:** The proposed approach is capable of segmenting various layers of DME subjects with large deformation. The approach has superior performance (0.89) in comparison with standard baseline (0.87 and 0.82) in terms of F-score. **Conclusion** A model can be trained with capability of predicting multiple layer edges with reduced in training time complexity and without trading off evaluation performance.

Keywords: Optical Coherence Tomography; Diabetic Macular Edema; Image Segmentation; Machine learning.

Ophthalmology involves diagnosis of eye affected systemic diseases along with eye pathologies. Conventionally selection of medical imaging instrument (fundus camera, fundus auto-fluorescence, etc.) by ophthalmologist is subjective to pathology. Adaptation of optical coherence tomography (OCT) a subsurface imaging technique for cross-sectional interpretation of eye has been a paradigm shift in eye imaging [1,2,3,4,5]. This led to attributing cross-sectional information of eye to various pathologies. Common clinical practice involve quantification of crossection information through layers width profiles[6,7,8,9,10]. To reduce the work load of ophthalmologists layer delineation pipeline has been automated through segmentation algorithms [11-19]. Pathologies related to retina are considered major reason for irreversible vision loss so majority of OCT segmentation algorithms aims at segmentation of various layers in retina. To illustrate the robustness of the proposed approach OCT scans of subjects with diabetic macular edema (DME) are considered for experimentation as it accounts for one of leading causes of blindness in the elderly. With the driven importance of retinal layer thickness quantification various image segmentation algorithms ranging from A-scan line (each column in image) based approaches to graph approaches including pattern recognition as intermediary step. Graph based approaches have illustrated resilience to noise and small layer deformations. But, the performance is subjective to the estimation of the edge maps. Major practices involve gradient derivatives (Eigen values) or classifiers (machine learning) for prediction of edge maps. Machine learning (ML) approaches have an advantage of modeling complex feature patterns for edge detection compared to heuristic rules [17,19,20,21]. Traditional ML based approaches involve prediction of a scalar value given a feature vector. Structured prediction is an evolving field in ML where the algorithm predicts the output structure given a feature vector[22,23]. In case of edge prediction given a feature vector at a location, an edge patch is predicted instead of a scalar value representing probability of edge presence at that corresponding location. Structured forests for edge detection (SFE) is a structured prediction algorithm which can be extended for detection of layer information along with edges. Incorporation of such information resulted in effective delineation of retinal layers particularly in case of subjects with large layer deformation. The layer information is incorporated through one vs. all approach where each model is trained for individual layer edges [23]. Such practice requires n models for n layers which

[1]School of Medical Science and Technology, IIT Kharagpur, Kharagpur, India, [2]Department of Mathematics, IIT Kharagpur, Kharagpur, India
Corresponding Author : SPK Karri, E-mail: phanimoresmiles
Received on : 30/10/2016, Revision accepted on : 10/12/2016
Conflict of Interest : None, Financial Disclosure : None

effectively increases the training time complexity. The proposed approach explores the possibility of modeling multiple layer edges by single model so the training time complexity is significantly reduced and helpful in case of delineating large number of layers. It is also required to achieve this without trading off the prediction error. The proposed model also retains the salient characteristics of traditional SFE i.e., resilient to noise, robust to large deformations, handling shadow artifacts, etc.). SFE modeling requires images corresponding segmented ground truths and edges along the object boundaries.

Given a set of images(m x n), SFE preprocessing step involves filtering images (m x n) for noise removal, feature quantification where histogram of oriented gradients with d-1 bins are extracted at each pixel, appending image with feature resulting in tuple (m x n x d), feature patches are extracted from tuple with each patch being 32 x 32 x d. Given segmented ground truths and edges corresponding to images, label patches and edge patches of size 16 x 16 are extracted retaining the location correspondence between feature patch central pixel, label patch central pixel and edge patch central pixel.

Similar to traditional random forests structured forests are an ensemble of decision trees. During training, at each tree‘s decision node label patches are projected onto a low dimensional space to create binary vectors, clustered binary vectors into two classes and decision rule is constructed (gini impurity) such that features patches takes up the class same as corresponding label patch‘s binary vector with minimum error. Each feature patch is driven to either left child path or right child path of a decision node based on the constructed decision rule of the node.

As each child path is appended with decision node the label patches orresponding to each feature patch in a child path are retrieved and decision rule construction followed by child path patch separation are implemented. The process is repeated for each child path until the similarity between label patches in the path fall under a threshold. In such cases child path is appended with leaf node. At each leaf node label patches and corresponding edge patches are retrieved and mean patches of each set (one for label patches and the other for edge patches). During testing given feature patch the decision rules directs the feature path to a leaf node and the mean edge patch is considers as prediction of the tree. Average of individual tree prediction is treated as forest prediction. Traditional SRF doesn’t care for the object class contributing to the edge. To include class information through ‘one vs. all’ approach involves creation of n (number of classes) sets of data where i^{th} set include feature patches, label patches treating only class ‘i’ as foreground and edge patches suppressing edge values not falling under i^{th} class object. Each dataset is trained with an independent model so during prediction the i^{th} model is expected to predict the i^{th} class edge. It is identified the bottle neck of the information is due to projecting the segmentation patches to low dimension for creating binary vectors. If scalar vectors (multiple values) are created instead of binary vectors results in construction of decision rules for object specific edges.

Material and method:

An online Duke OCT dataset comprising DME subject’s retinal scans is considered for experimentation. The dataset also includes contour indexes for eight layers along each column of image are annotated by two experts, heuristically edges based graph segmentation and kernel regression feature based graph segmentation. The algorithms requires retinal oct images (during training), corresponding segmentation ground truths (label) of each layer and upper boundary (edges) of each layer with unique value. The algorithm being a machine learning based approach it involves three phases: data preparation, training and prediction. Data preparation involves quantification of HOG features at each pixel of an OCT image, concatenated along the depth resulting in a tuple, tuple patch creation with same depth as tuple and finally patch extraction of label image and edge image retaining the central pixel correspondence with tuple edge. Training process involve constructing of decorrelated trees where each tree is trained with different data and decision rules at each decision node are identified based on tuple patches and corresponding label patches. It also includes the mean label patch and edge patch in each leaf node of each tree. Prediction process employs a test image tuple patch and identifies a leaf node so the mean edge stored in the leaf node is treated as prediction.

Graph based retinal layer segmentation approaches involve identification edges at each pixel based on heuristics or machine learning and recent approaches are more focused on machine learning approaches due to the performance. Machine learning approaches don’t predict edges inherently as they are trained for segmentation and edges are computed from individual layer probability maps.

For each layer, edge (upper contour) map of corresponding layer is considered, column of values '1' is appended on left and right hand side of the edge map, a dynamic program is employed to identify shortest path from top left pixel to right bottom pixel of the image, such path is smoothened with a filter and outcome is treated as elucidation between considered layer and upper layer.

During training traditional random forest decision rule construction involves minimization of information loss between feature vectors and corresponding class labels. Structured random forest need to consider label patch instead of class label (a scalar) is computationally expensive. So label patches are projected on to low dimensional space through random sampling instead of conventional approaches like principle component analysis etc. Such approach is adopted with the idea for introduction of noise during decision rule construction which has been empirically proven to resolve overfitting problem. The low dimensional projection being a random sampling replacing binary vector the proposed idea of replacing binary vector with scalar vector will not burden the projection process. During clustering instead of computing Euclidean distance between two vectors with two states the distance between two vectors with 'n+1' states is computed. Even this will not impose any computational overhead because the scalar vectors is of data type unidentified integer but not float. From algorithm point alterations in few lines of structured random forest source code is sufficient for constructing low dimensional scalar vector during training.

The label and edge patches are averaged and stored in leaf nodes. The mean patches takes multiple states rather to anticipated due to low space projection and averaging results. This results in artifacts model prediction. To avoid this an one time refinement step is introduced after training where 'n' models are created out of trained single model so i^{th} model predicts the edge patches corresponding to the i^{th} layer. This is achieved by clipping of the mean edge patch values with i+1 for i^{th} model. As stated before given a test image, feature tuple is generated, tuple patches retaining the depth with a stride of 1 are generated, i^{th} model indexes each tuple patch to one of mean edge patches and finally overlapping regions between edge patches are averaged to reconstruct the edge map of i^{th} layer with dimensions same as test image.

Results:

OCT scans of 10 patients were included in the experimentation. A total of 55 images from first five subjects are considered for training and remaining 55 images are considered as testing set Multiscale information is embedded by downscaling and upscaling retinal images and concatenate the extracted HOG features to original image tuple along depth. During training, generation of tuple patches with stride '1' results in abundant metadata and an overload on model training. So a 1 million patches along the contours within 8px radius are considered and an additional 1.5 million patches far from contours are considered. The clustering at each decision node is performed by employing PCA and gini measure is considered for decision rule construction. The random forest with six trees is trained. During testing for each image instead of extracting tuple patches with at stride '2' the extraction is performed at stride '2' to reduce the prediction time complexity. Training of proposed approach takes approximately 14 hrs for learning layers of all edges where as one vs. all approach takes around 12 hrs per layer edge resulting few days for all layers. To evaluate the prediction

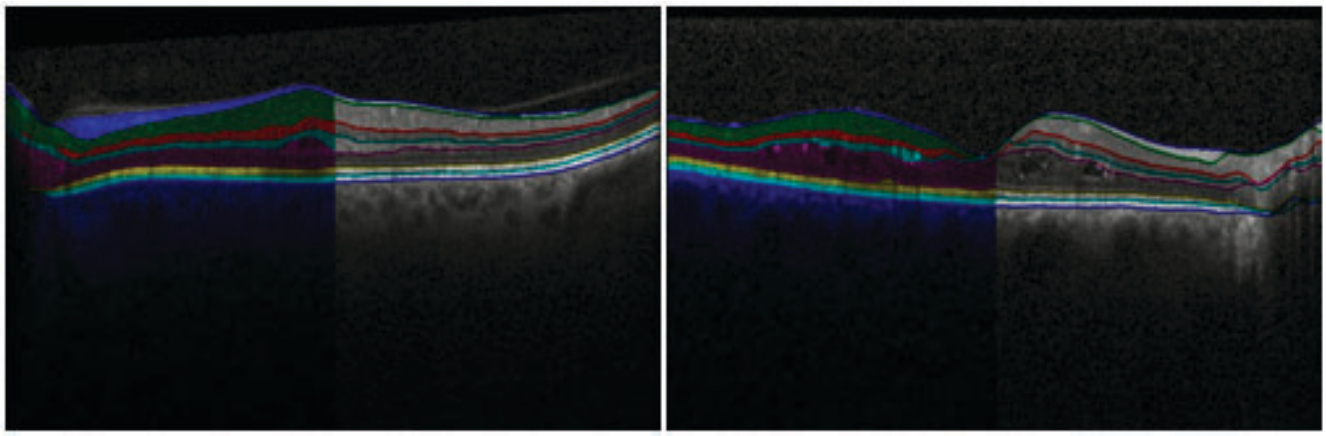

Figure 1: Illustrating the segmentation capability of proposed methods in normal images and pathological images.

capability the segmentation performance is compared with F-score metric. To illustrate that the model retains the salient points of one Vs all approach the segmentation results in cases of noise, large deformation and shadow artifacts are shown in figure below.

An approach for training single model capable of predicting layer specific edges has been proposed. The one vs. all approach has been proposed to validate a hypothesis that object or layer specific edges could be learned. To speed up the training multi-boundary (in feature space) approach or single model is employed which requires constructing complex functions. the random forests are well established for handling such complex boundaries and Table 1 supports the same based on prediction error. It is evident that time taken for training a single model is less to training 'n' models each corresponding to a layer. The prediction time complexity is same as one vs. all approach because 'n' edge maps (each for a layer) need to be generated due to the employment of dynamic programming. As to our knowledge dynamic program cannot identify multiple shortest paths simultaneously subjective to edges from different topologies (layers).

Table 1: F-score of proposed method compared to benchmarks (AN and AD) and comparable performance against 'One Vs all' approach with large training complexity.

	Proposed	AN[17]	AD[17]	One Vs all[23]
ILM	0.87	0.778	0.853	0.87
NFL/GCL	0.91	0.772	0.895	0.90
IPL/INL	0.81	0.652	0.757	0.80
INL/OPL	0.76	0.67	0.747	0.77
OPL/ONL	0.94	0.868	0.931	0.94
ISM	0.88	0.878	0.872	0.88
RPE	0.86	0.823	0.824	0.86

Incorporation of label and edge patches with pixels from multiple states instead of binary states has impacted functionally during training in low dimensional projection for scalar vector, clustering and mean patch values at leaf nodes. Computationally both one vs. all and single model approach are the same during training and testing i.e., computational complexity of training is higher to testing. The proposed approach still retinas the parallizable aspect of the one vs. all approach i.e., during training data can be segregated into six sets (equal to number of trees) each with 25\% (user defined) of the data and each tree can be trained mutually exclusive. During testing as 'n' models are created for one time refinement step and each model has six trees where each tree can be evaluated independently.

Machine learning based approaches majorly influenced by data samples, features quantified and parameter settings. Data samples are more effective if they represent the majority of the population so providing more data which is self replicative constrains the generalizability of the model. This is the reason for choosing more number of non-edge patches in comparison to edge patches as non-edge patches have more patterns to comprehend. Such imbalance could bias the prediction of the algorithm but random forests algorithm is proven to be resilient to data imbalance. It is observed that increasing the ratio improves the performance but burdens the computations. It is also noted that increasing cell size during HOG feature quantification instead of patch size has larger impact compared to computational burden it put on. There is no quantifiable approach for selection of number of trees other to training models with different number of trees and identifying the break point for overfitting through test error.

The approach can be extended to any medical based low vision problem involving edge detection for classification or object specific enhancement. The proposed approach aims at reinforcement of dynamic programming. The approach can be extended to any medical based low vision problem involving edge detection for classification or object specific enhancement. The proposed approach aims at reinforcement of dynamic programming based segmentation methods but not to be mis-categorized with segmentation approaches for retinal layers. As edge is a common factor for multiple graphs based approaches this can be appended as preprocessing step for any off the shelf graph based segmentation. The trained model requires less space and computational complexity so it can be deployed to remote clinics. On inflow new data instead of retraining entire model a set of trees can be trained and can be merged with existing trees on prediction end or prediction end based on the feasibility.

The proposed approach is aimed at reducing the time complexity of training without trading off the prediction capability in comparison to one vs. all approach. It is

capable of handling various adverse but fails in case of low gradient conditions. The OCT scans being 3D in nature future work involve incorporation of neighboring scans information to make the model more resilient to low gradients. As consideration of HOG features being the bottle neck of information processing incorporation of deep learnt features could be another path to be explored to handle low gradients.

Discussion

The proposed algorithm is able to expand the capability of standard edge forests for modeling multiple layer specific edges with single random forest. It is also established that such modeling did not tradeoff the perdition error for the application of edge based retinal layer segmentation. The proposed algorithms has similar space and computational complexity in training phase and prediction phase in comparison to one vs. all approach. The algorithm is not capable of predicting edges under low gradients as features are responsible.

References:

1. Nassif N, Cense B, Park B, Pierce M, Yun S, Bouma B, Tearney G, Chen T, de Boer J. In vivo high-resolution video-rate spectral-domain optical coherence tomography of the human retina and optic nerve. Optics express 2004;12: 367-76.
2. Jeoung JW, Park KH, Kim TW, Khwarg SI, Kim DM. Diagnostic ability of optical coherence tomography with a normative database to detect localized retinal nerve fiber layer defects. Ophthalmology 2005;112: 2157-63.
3. Medeiros FA, Zangwill LM, Bowd C, Vessani RM, Susanna R, Weinreb RN. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. American journal of ophthalmology. 2005;13: 44-55.
4. Na JH, Sung KR, Baek S, Kim YJ, Durbin MK, Lee HJ, Kim HK, Sohn YH. Detection of Glaucoma Progression by Assessment of Segmented Macular Thickness Data Obtained Using Spectral Domain Optical Coherence TomographySegmented Macular Layer Thinning in Glaucoma. Investigative ophthalmology & visual science. 2012; 53: 3817-26..
5. Farsiu S, Chiu SJ, O'Connell RV, Folgar FA, Yuan E, Izatt JA, Toth CA, Age-Related Eye Disease Study 2 Ancillary Spectral Domain Optical Coherence Tomography Study Group. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. Ophthalmology 2014; 121: 162-72.
6. Mayer MA, Hornegger J, Mardin CY, Tornow RP. Retinal nerve fiber layer segmentation on FD-OCT scans of normal subjects and glaucoma patients. Biomedical optics express. 2010; 11:1358-83.
7. Hood DC, Raza AS, Kay KY, Sandler SF, Xin D, Ritch R, Liebmann JM. A comparison of retinal nerve fiber layer (RNFL) thickness obtained with frequency and time domain optical coherence tomography (OCT). Optics express. 2009; 17: 3997-4003.
8. Kanamori A, Nakamura M, Escano MF, Seya R, Maeda H, Negi A. Evaluation of the glaucomatous damage on retinal nerve fiber layer thickness measured by optical coherence tomography. American journal of ophthalmology. 2003; 135: 513-20.
9. Drexler W, Fujimoto JG. State-of-the-art retinal optical coherence tomography. Progress in retinal and eye research. 2008; 27: 45-88.
10. Wu Z, Huang J, Dustin L, Sadda SR. Signal strength is an important determinant of accuracy of nerve fiber layer thickness measurement by optical coherence tomography. Journal of glaucoma. 2009;18: 213-6.
11. Tátrai E, Ranganathan S, Ferencz M, DeBuc DC, Somfai GM. Comparison of retinal thickness by Fourier-domain optical coherence tomography and OCT retinal image analysis software segmentation analysis derived from Stratus optical coherence tomography images. Journal of biomedical optics 2011; 16:056004.
12. Kafieh R, Rabbani H, Abramoff MD, Sonka M. Intra-retinal layer segmentation of 3D optical coherence tomography using coarse grained diffusion map. Medical image analysis 2013;17: 907-28.
13. Drexler W, Liu M, Kumar A, Kamali T, Unterhuber A, Leitgeb RA. Optical coherence tomography today: speed, contrast, and multimodality. Journal of biomedical optics. 2014; 19: 071412.
14. Chiu SJ, Li XT, Nicholas P, Toth CA, Izatt JA, Farsiu S. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. Optics express.2010; 18:19413-28.
15. Mwanza JC, Oakley JD, Budenz DL, Chang RT, O'Rese JK, Feuer WJ. Macular ganglion cell–inner plexiform layer: automated detection and thickness

reproducibility with spectral domain–optical coherence tomography in glaucoma. Invest Ophthalmol Vis Sci.2011; 152: 8323-9.

16. Srinivasan PP, Heflin SJ, Izatt JA, Arshavsky VY, Farsiu S. Automatic segmentation of up to ten layer boundaries in SD-OCT images of the mouse retina with and without missing layers due to pathology. Biomedical optics express 2014; 15:348-65.
17. Chiu SJ, Allingham MJ, Mettu PS, Cousins SW, Izatt JA, Farsiu S. Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. Biomedical optics express 2015; 16:1172-94.
18. Dufour PA, Ceklic L, Abdillahi H, Schroder S, De Dzanet S, Wolf-Schnurrbusch U, Kowal J. Graph-based multi-surface segmentation of OCT data using trained hard and soft constraints. IEEE transactions on medical imaging 2013; 32: 531-43.
19. Rathke F, Schmidt S, Schnörr C. Probabilistic intra-retinal layer segmentation in 3-D OCT images using global shape regularization. Medical image analysis. 2014;18: 781-94.
20. Prasad M, Zisserman A, Fitzgibbon A, Kumar MP, Torr PH. Learning class-specific edges for object detection and segmentation. In Computer Vision, Graphics and Image Processing 2006 (pp. 94-105). Springer Berlin Heidelberg.
21. Lucchi A, Li Y, Smith K, Fua P. Structured image segmentation using kernelized features. In Computer Vision–ECCV 2012 (pp. 400-413). Springer Berlin Heidelberg.
22. Dollár P, Zitnick CL. Structured forests for fast edge detection. In Proceedings of the IEEE International Conference on Computer Vision 2013 (pp. 1841-1848).
23. Karri SP, Chakraborty D, Chatterjee J. Learning layer-specific edges for segmenting retinal layers with large deformations. Biomedical Optics Express 2016; 17:2888-901.

Cite this article as:
Karri SPK, Chakraborty D, Chatterjee J. Rapid training of layer specific edges for segmentation of retinal layers in OCT. Current Indian Eye Research 2016; 3:73-8.