



# Deep Random Walk for Drusen Segmentation from Fundus Images

Fang Yan<sup>1</sup>, Jia Cui<sup>1</sup>, Yu Wang<sup>1</sup>, Hong Liu<sup>1,2</sup>, Hui Liu<sup>3</sup>, Benzheng Wei<sup>4</sup>,  
Yilong Yin<sup>5</sup>, and Yuanjie Zheng<sup>1,2,6,7</sup>(✉)

<sup>1</sup> School of Information Science and Engineering, Shandong Normal University,  
Jinan, China

yjzheng@sdsnu.edu.cn

<sup>2</sup> Shandong Provincial Key Lab for Distributed Computer Software Novel  
Technology, Jinan, China

<sup>3</sup> Department of Biomedical Engineer, Dalian University of Technology,  
Dalian, China

<sup>4</sup> College of Science and Technology, Shandong University of Traditional Chinese  
Medicine, Jinan, China

<sup>5</sup> School of Computer Science and Technology, Shandong University, Jinan, China

<sup>6</sup> Institute of Biomedical Sciences, Shandong Normal University, Jinan, China

<sup>7</sup> Key Lab of Intelligent Computing and Information Security in Universities  
of Shandong, Jinan, China

**Abstract.** This paper presents a deep random walk technique for drusen segmentation from fundus images. It is formulated as a deep learning architecture which learns deep representations from fundus images and specify an optimal pixel-pixel affinity. Specifically, the proposed architecture is mainly composed of three parts: a deep feature extraction module to learn both semantic-level and low-level representation of image, an affinity learning module to get pixel-pixel affinities for formulating the transition matrix of random walk and a random walk module which propagates manual labels. The power of our technique comes from the fact that the learning procedures for deep image representations and pixel-pixel affinities are driven by the random walk process. The accuracy of our proposed algorithm surpasses state-of-the-art drusen segmentation techniques as validated on the public STARE and DRIVE databases.

**Keywords:** Drusen segmentation · Retinal fundus images  
Deep feature extraction · Affinity learning · Random walk

## 1 Introduction

Drusen are a kind of degenerative disease that occurs in choroidal retina. They are caused by abnormal deposition of metabolites from retinal pigment epithelium(RPE) cells. Moreover, drusen are the main manifestations of age-related macular degeneration(AMD), at the same time they are also the major causes of blindness in the elderly [1]. Longitudinal studies show that eyes with larger size

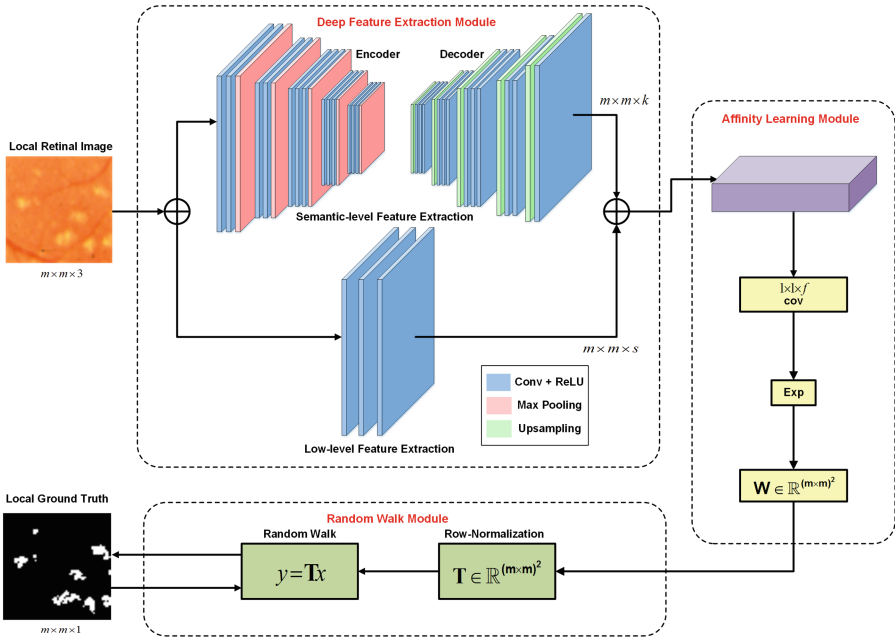
or number of drusen are more likely to cause degeneration of pigment epithelial cells, leading to a decline of central vision [2,3]. Therefore, the evaluation of the areas, locations and quantity of drusen from retinal fundus images is of great significance in the clinic especially in the remote screening and diagnosis of AMD.

The main challenges for drusen segmentation lie in three factors that are color or brightness, shape and boundary fuzziness of drusen. For color or brightness, drusen in yellowish-white which is close to the color of fundus image and optic disc. Moreover, drusen also present the characteristic of uneven brightness and the interference of factors such as blood vessels, which have a great impact on the accuracy of the segmentation. In the aspect of shape, drusen often present irregular shapes or circles, and have obvious changes in size, scattering in the vascular arch. For the boundary fuzziness, there is no obvious boundary for soft drusen, which increase difficulty for segmentation accuracy [4–6]. The deep feature extraction module used in this paper can effectively improve the accuracy in view of the semantic features and the low-level features.

There are a variety of drusen segmentation technologies in the field of ophthalmology image research. In this paper, we mainly extract semantic features from fundus images based on the characteristics of drusen, and then acquire classified labels via random walk to detect the locations and areas of drusen. There are various techniques for drusen segmentation approaches, for instance, in frequency domain [7], thresholding methods [8], and feature extraction [3]. Specially, in the aspect of feature extraction, many features have been used in the previous work, such as the image gray value [6,9], Hessian features and intensity histogram, total variation features [5], etc. For deep learning semantic segmentation methods, lots of networks have tried in improving accuracy, like traditional method [10], fully convolutional networks (FCNs) [11], deep convolutional nets (SegNet) [12], multi-path refinement networks (RefineNet) [13], etc. Though most of the existing methods can be used as good references in drusen segmentation, there are still some restrictive problems. First, most drusen segmentation methods still use manual feature and cannot get deeper and lower-level information. Second, semantic segmentation is easy to make results smooth. Third, this kind of methods are rarely applied in drusen segmentation now.

In this article, we propose a novel deep random walk network for the implementation of drusen segmentation from retinal fundus images. It extracts the semantic-level and low-level features of patches which are generated from fundus images as training data to the net, and then constructs a translation matrix storing pixel-pixel affinities. Inspired by random walk methods, the framework constructs an end-to-end training network combining the stochastic initial status of the input image with pixel-pixel affinities. Specifically, we obtain the feature maps across an encoder-decoder structure and a refined fully convolutional network, and the whole structures can be jointly optimized. Therefore, the progress not only reduces the parameters, but also preserves the edges information under the condition that the spatial information is not lost, which finally improves the accuracy.

The proposed method can effectively solve challenges above in dealing with drusen segmentation problems, and specific advantages are as follows. Firstly, compared with traditional approaches extracting manual features, we combine semantic information representations with low-level feature extraction method which makes up for the lack of edge smoothing in the process of semantic feature extraction. This is crucial to drusen photos because of characteristics of images themselves. Secondly, the application of random walk approach which is matrix multiplication in mathematics can guarantee the implementation of back propagation algorithm in training process. Finally, the integration of feature descriptions, pixel-level affinities learning and random walk to do classification can be jointly optimized to form an end-to-end network. This results in the dimensionality of parameters space reduced. Based on above advantages, the experiments also prove that our method can improve the accuracy of drusen segmentation.



**Fig. 1.** The architecture of our proposed drusen segmentation method. Deep random walk networks contain 3 modules for deep feature extraction, affinity learning and random walk for drusen segmentation, respectively.

## 2 Deep Random Walk Networks

The proposed deep random walk networks aim to detect and segment locations and areas of drusen from retinal fundus images. Given color fundus images and corresponding ground truth as training materials, we divide them into patches

whose size is  $m \times m$  in order to solve the problem of fewer medical samples. When selecting training data,  $n$  patches were sampled stochastically from drusen and non-drusen regions. We represent the training data as  $\{S_1, S_2, \dots, S_n\}$ , and  $n$  denotes the number of training images.

Three main modules of deep random walk architecture were integrated to extract both semantic-level and low-level features and construct transition matrix which represents relationship between pixels. Figure 1 shows a schematic illustration of our framework. Deep feature extraction module aims at semantic and low level information’s extraction. Affinity learning Module formulates the transition matrix of random walk. And the random walk module aims to acquire manual labels. Random walk is a form of matrix multiplication in mathematics. This form helps to optimize the three modules in the network and achieves end-to-end training process using the stochastic gradient descent. The detailed description is as follows.

## 2.1 Deep Feature Extraction Module

The feature extraction module consists of two branches, a semantic-level feature extraction branch which learns deep information based semantic features and a low-level feature extraction branch which acquires detailed features such as sharper edges to improve accuracy. Then the obtained descriptions of image features can be used to represent pixel-pixel affinities in affinity learning module.

For semantic-level feature extraction branch, we get dense feature maps through a encoder and decoder network corresponding to SegNet [12] which considers fundus image patches as training input. Different from SegNet, we obtain feature maps via encoder and decoder network and put the dense representations to affinity learning module to acquire relationships between pixels and then detect drusen in random walk module instead of soft-max classifier. The encoder network transforming the input fundus patches to downsampled feature representations is identical to the VGG16 network [14]. Moreover, it is composed by 13 convolutional layers related to the first 13 convolutional layers in VGG16 network and 5 max pooling layers which carry out with  $2 \times 2$  windows and the stride is 2. Specially, an element-wise rectified-linear non-linearity(ReLU)  $\max(0, x)$  is applied before the max pooling layer. The decoder network upsamples the feature maps learnt from the encoder network using upsampling layers and convolutional layers. In order to reduce the loss of spatial resolution, it is needful to transform max pooling indices from encoder network to upsampling layers in decoder network.

For low-level feature extraction branch, it consists of 3 convolutional layers, each of which followed by non-linear “ReLU” layers. The goal of this branch is to acquire low-level information such as sharper edges missed in front branch for the encoder and decoder networks sometimes result in overly smooth. The detailed illustration is shown in [15].

Compared to the structure of [15], the semantic-level network is in parallel action with low-level network instead of the concatenation. The output of

semantic-level network is  $m \times m \times k$ , where  $m$  denotes the length of the input square patch and  $k$  represents the number of features. Similarly, the output of low-level network is  $m \times m \times s$ , where  $s$  is the number of feature maps.

## 2.2 Affinity Learning Module

The target of the affinity learning module is to construct a transformation matrix which learns the information of pairwise pixel-pixel and is required in the random walk module. According to the semantic-level and low-level features obtained from the feature extraction module, we integrate the two feature maps into a matrix denoted as  $m \times m \times (k + s)$ . Then a new weight matrix ( $N_n \times f$ ) is generated via computing relationships between neighboring-pixel pairs, where  $N_n$  represents the total number of neighboring affinities and  $f$  is equal to  $2(k + s)$ . The neighborhood can be defined as 4-connection in this paper.

The affinity learning module consists of a  $1 \times 1 \times f$  convolutional layer and an exponential layer which normalize the obtained matrix  $W$ . Moreover, matrix  $W$  will be a sparse weight matrix after transformation to  $m^2 \times m^2$  and via a limited threshold computing in order to reduce the complexity.

## 2.3 Random Walk Module

Random walk can be expressed as a form of  $y = T \times x$ , where  $T$  storing the weight of pixel-pixel affinities is called the transformation matrix denoted as  $m^2 \times m^2$  via row-normalization of  $W$ , and  $x$  is the initial state recorded as  $m^2 \times 1$ . Here we can understand each pixel in the segmented image as a node in the space, and the relationships between each pair of nodes can be represented by weight values. This work, we take the initial value of  $x$  via the given initial segmentation using [6], and get the final stable potential via matrix multiplication. Finally, the segmented image are obtained via the softmax layer [16].

During testing, random walk algorithm converts the initial potential energy of image segmentation to the final potential energy via iterations. Furthermore, the terminational condition is that the energy of the image tends to be stable, which is to say the vector  $x$  is no longer changing. A detailed proof and deduction are presented in [16].

## 3 Implementation

We implemented the deep random walk network using Caffe, and carried training and testing process on a NVIDIA GTX 1080Ti graphics card. When in the training time, the fixed learning rate was 0.001 and the momentum was 0.9. The overall training phase was 30 epochs.

## 4 Experiment

### 4.1 Dataset

We evaluated the proposed deep architecture in two public datasets: STARE and DRIVE. The STARE dataset contains of 400 retinal fundus images, each of which has a size of  $700 \times 605$ . We selected 46 images containing drusen from 63 diseased images to verify our ideas, where we used 20 images for training and 26 images for testing. The DRIVE dataset includes 40 retinal fundus images, and each of them is  $768 \times 584$ . 9 photos were chosen to test our network. As shown in [3], the ground truth is marked manually via the computer drawing tool.

We make the number increase to 28,800 after augmentation by applying 18 rotations, 16 stretching effects and 5 bias fields on training images. In addition, we train our network with patches extracted from the training images, for which the number is nearly 2,880,000 taking patches with each size of which is  $64 \times 64$ . It is worth noting that it is allowed to be covered when selecting training patches in the same one eye image. In the prediction stage, we use a sliding window of  $64 \times 64$  to take tiles, and the stride is 64. Therefore, this is a non-covered selection.

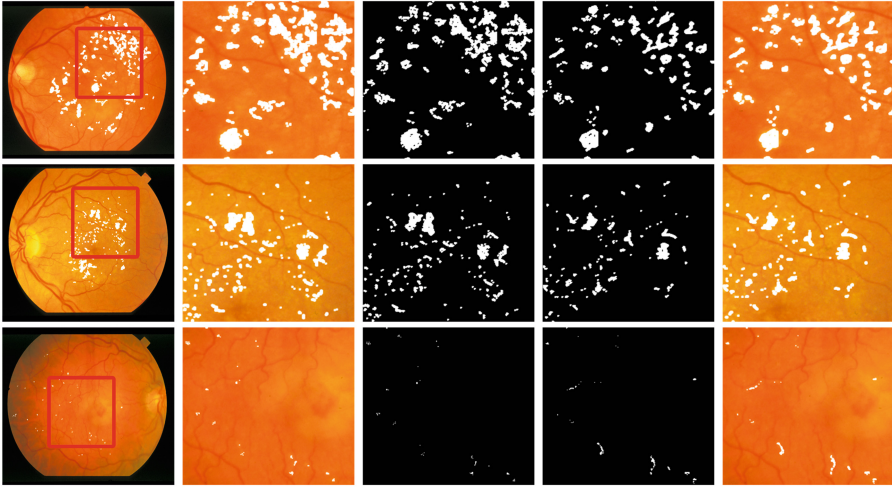
**Table 1.** Accuracy evaluation in dataset STARE and DRIVE.

Method	%Se	%Spe	%Acc
Our method	92.02	97.30	97.13
HALT [17]	90.13	94.37	94.89
Liu <i>et al.</i> [18]	84.00	87.94	88.38
Ren <i>et al.</i> [3]	90.12	96.84	96.56
Zheng <i>et al.</i> [6]	87.49	89.82	90.27

### 4.2 Evaluation and Result

The common evaluation methods of drusen segmentation are sensitivity(Se), specificity(Spe), and accuracy(Acc) [19], where Se refers to the rate of true positive detection, Spe represents the rate of false positive detection, and Acc measures the ration of the total correctly identified pixels [3,5]. According to the three evaluation indexes, we run our algorithm on public dataset STARE and DRIVE, and compare the results with four classical drusen segmentation approaches which are HALT [17], Liu *et al.* [18], Ren *et al.* [3] and Zheng *et al.* [6]. As the results shown in Table 1, our network can resolve the challenges of drusen segmentation better than other state-of-the-art techniques because the learned deep features can help to deal with color similarity of drusen to other tissues and drusen variations in shape and size. Moreover, the random walk process achieves precise segmentation at fuzzy drusen boundaries.

Figure 2 shows the segmentation results on three classical photos from the STARE dataset with large drusen, vague small and large drusen, and small sparse drusen. The results of drusen segmentation are satisfying because of the areas and locations are successfully detected. Our algorithm acquired satisfied segmented results due to the deep random walk network which extracts the semantic-level and low-level features.



**Fig. 2.** Segmentation results of different drusen types from the STARE dataset. From left to right: the entire original fundus image, drusen region in the retinal color image, ground truth, segmented image, result of our algorithm.

## 5 Conclusion

In this work, we introduced a deep random walk network for drusen segmentation from fundus images. Our technique formulated as a deep learning architecture extracts the semantic-level and low-level feature maps to construct pixel-pixel affinities. Inspired by the random walk method, our structure constructs an end-to-end training network and the accuracy of our proposed algorithm surpasses state-of-the-art drusen segmentation techniques. Our future work would include experimenting with other frameworks in order to alternate to the deep random walk network. In addition, we would like to extend our net to other domains such as matting and so on.

**Acknowledgements.** This work was made possible through support from Natural Science Foundation of China (NSFC) (61572300) and Taishan Scholar Program of Shandong Province in China (TSHW201502038).

## References

1. Brandon, L., Hoover, A.: Drusen detection in a retinal image using multi-level analysis. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2878, pp. 618–625. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-39899-8\\_76](https://doi.org/10.1007/978-3-540-39899-8_76)
2. Sarks, S.H., Arnold, J.J., Killingsworth, M.C., Sarks, J.P.: Early drusen formation in the normal and aging eye and their relation to age related maculopathy: a clinicopathological study. *Br. J. Ophthalmol.* **83**(3), 358–368 (1999)
3. Ren, X., et al.: Drusen segmentation from retinal images via supervised feature learning. *IEEE Access* **PP**(99), 1 (2017)
4. Schlanitz, F.G., et al.: Performance of drusen detection by spectral-domain optical coherence tomography. *Investig. Ophthalmol. Vis. Sci.* **51**(12), 6715 (2010)
5. Zheng, Y., Wang, H., Wu, J., Gao, J.: Multiscale analysis revisited: detection of drusen and vessel in digital retinal images. In: *IEEE International Symposium on Biomedical Imaging: From Nano To Macro*, pp. 689–692 (2011)
6. Zheng, Y., Vanderbeek, B., Daniel, E., Stambolian, D.: An automated drusen detection system for classifying age-related macular degeneration with color fundus photographs. In: *IEEE International Symposium on Biomedical Imaging*, pp. 1448–1451 (2013)
7. Barriga, E.S., et al.: Multi-scale am-fm for lesion phenotyping on age-related macular degeneration. In: *IEEE International Symposium on Computer-Based Medical Systems*, pp. 1–5 (2009)
8. Shin, D.S., Javornik, N.B., Berger, J.W.: Computer-assisted, interactive fundus image processing for macular drusen quantitation. *Ophthalmology* **106**(6), 1119–25 (1999)
9. Smith, R.T.: Automated detection of macular drusen using geometric background leveling and threshold selection. *Arch. Ophthalmol.* **123**(2), 200–206 (2005)
10. Shotton, J., et al.: Real-time human pose recognition in parts from single depth images. *Commun. ACM* **56**(1), 1297–1304 (2011)
11. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015)
12. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **PP**(99), 1 (2017)
13. Lin, G., Milan, A., Shen, C., Reid, I.D.: Refinenet: multi-path refinement networks for high-resolution semantic segmentation. In: *CVPR*, vol. 1, no. 2, 5 p. (2017)
14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *Comput. Sci.* arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)
15. Xu, N., Price, B., Cohen, S., Huang, T.: Deep image matting. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 311–320 (2017)
16. Bertasius, G., Torresani, L., Yu, S.X., Shi, J.: Convolutional random walk networks for semantic image segmentation, pp. 6137–6145 (2016)
17. Rapantzikos, K., Zervakis, M., Balas, K.: Detection and segmentation of drusen deposits on human retina: potential in the diagnosis of age-related macular degeneration. *Med. Image Anal.* **7**(1), 95–108 (2003)
18. Liu, H., Xu, Y., Wong, D.W.K., Liu, J.: Effective drusen segmentation from fundus images for age-related macular degeneration screening. In: *Asian Conference on Computer Vision*, pp. 483–498 (2014)
19. Briggs, D.A.H.: Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* **17**(5), 479 (2000)