A coarse-to-fine deep learning framework for optic disc segmentation in fundus images

Lei Wang [a], Han Liu [a], Yaling Lu [b], Hang Chen [c], Jian Zhang [c,*], Jiantao Pu [a,*]

[a] Departments of Radiology and Bioengineering, University of Pittsburgh, Pittsburgh, PA 15213, USA
[b] Department of Electrical and Electronic Engineering, Wuhan Polytechnic University, Wuhan, 430023, China
[c] Department of Ophthalmology, Shaanxi Provincial People's Hospital, Xi'an, 710068, China

ARTICLE INFO

ABSTRACT

Accurate segmentation of the optic disc (OD) depicted on color fundus images may aid in the early detection and quantitative diagnosis of retinal diseases, such as glaucoma and optic atrophy. In this study, we proposed a coarse-to-fine deep learning framework on the basis of a classical convolutional neural network (CNN), known as the U-net model, to accurately identify the optic disc. This network was trained separately on color fundus images and their grayscale vessel density maps, leading to two different segmentation results from the entire image. We combined the results using an overlap strategy to identify a local image patch (disc candidate region), which was then fed into the U-net model for further segmentation. Our experiments demonstrated that the developed framework achieved an average intersection over union (IoU) and a dice similarity coefficient (DSC) of 89.1% and 93.9%, respectively, based on 2978 test images from our collected dataset and six public datasets, as compared to 87.4% and 92.5% obtained by only using the sole U-net model. The comparison with available approaches demonstrated a reliable and relatively high performance of the proposed deep learning framework in automated OD segmentation.

1. Introduction

As one of the fundamental anatomic structures of the retina, the optic disc (OD) [1–3] appears as a bright circular region on color fundus images. The OD is the location of the optic nerve and the entry point for major blood vessels [4,5]. This unique anatomy makes the OD a landmark for disease assessment [6–8]. Although there is relatively high contrast between the OD and the surrounding regions on color fundus images, identifying the OD is often a challenging task [9,10], because the image quality can be confounded by a number of factors, such as light exposure, camera focus, motion artifact, and existing diseases. These factors may lead to inhomogeneous image quality [11,12], as illustrated by the examples in Fig. 1, where the OD locations and their boundaries are not well depicted.

In the past, extensive efforts [13–18] have been made to develop computerized methods for identifying the OD depicted on fundus images. Most of these methods [19,20] leveraged the anatomical characteristics of the OD and its relationship with other structures or anatomical landmarks, such as the elliptical / circular shape of the OD [21–23] and the parabola shape of the blood vessels [24,25], to initially locate the OD. On the basis of the OD location, additional procedures were used to identify the OD's boundaries [7]. The experiments showed that these methods had limited segmentation accuracy when facing fundus images with weak contrast and / or severe inhomogeneity. In this study, we proposed to approach this segmentation issue using deep convolutional neural network (CNN) technology [26–28]. The most distinctive advantage of this technology [29,31,32] is the ability to precisely learn and capture a large number of underlying image features with varying hierarchies and locations, and to optimally integrate these features to obtain a desirable result. Considering the unique characteristics of retinal images, we developed a novel coarse-to-fine deep learning framework on the basis of the U-net model [29]. Specifically, we firstly extracted retinal blood vessels from color fundus images and then constructed a vessel density map [30]. The fundus images and their vessel density maps were separately fed into the given U-net model to obtain two different segmentation results, which were combined further to obtain the possible location of the OD. Segmentation of the OD was finally performed using the same U-net model. A detailed description of this method and its performance validation follows.

* Corresponding authors.
   E-mail addresses: ykzhang_123@163.com (J. Zhang), puj@upmc.edu (J. Pu).
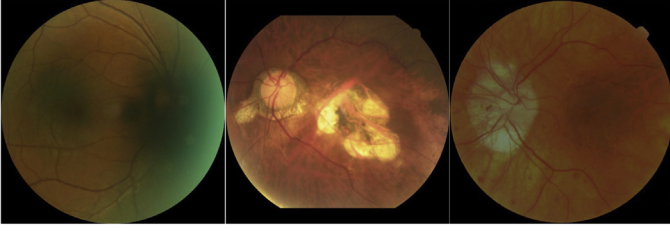
Fig. 1. Illustration of three fundus images with severe inhomogeneity and anomalous tissue variability.

## 2. Methods and materials

2.1. Scheme overview

The architecture of the developed method for segmenting the optic disc in a coarse-to-fine way consists of several components, as shown in Fig. 2. We derived the vessel density map from the retinal vessels to highlight the location of the optic disc with respect to the vessels. We used the U-net model to segment the OD depicted on color fundus images and their vessel density map, respectively. This resulted in two different segmentations that often contain multiple OD candidate regions. Thereafter, we used an overlapping strategy to exclude the false segmentation and identify the candidate region with the highest probability of being the OD. Applying the U-net model to the final candidate regions led to the identification of the optimal OD region.
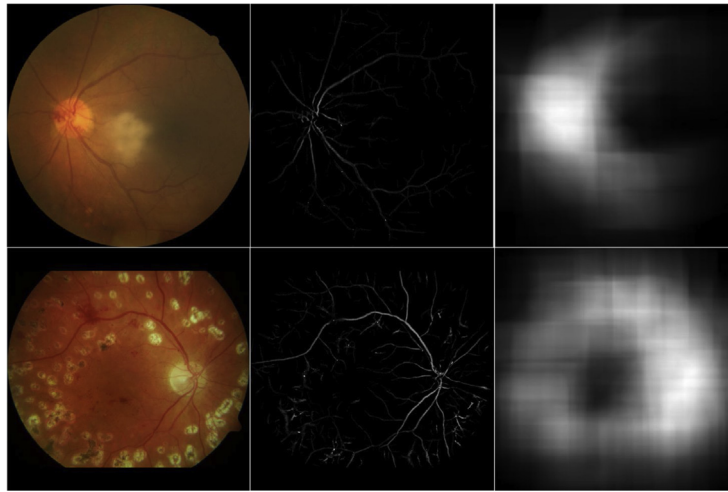
2.2. Vessel density map

The OD is the entry point for the retinal blood vessels. From the branching patterns of the retinal vessels, it is relatively straightforward for our human eyes, but not for the computer, to estimate the location of the OD. In this study, we proposed to use the vessel density map [30] to characterize the underlying property of the retinal vessels and their relationship with the OD. First, we used the method presented in [33] to extract the retinal vessels, which was represented as a grayscale image. On the retinal vessel images, each pixel denotes the contrast between the vessels and the background, and has large intensity for major vessels near the disc region. Then, we computed the vessel density map using the formulation:

$$D(x) = \sum_{y \in \Omega_x} V(y) \tag{1}$$

where $V(\cdot)$ and $D(\cdot)$ denote the pixel value of the vessel image and its density map at a given position, respectively. $\Omega_x = \{y : |x - y| \leq R\}$ is a given local region centered at x, with a radius of R. $D(\cdot)$ is the sum of the intensities of pixels in vessel feature image and actually denotes the density of the vessels within the region Ω_x. This strategy is different from the map proposed in paper [30], which counted the number of vessel pixels in local regions. In contrast, our strategy considered the density values of the vessel pixels. The difference enables the proposed density to emphasize the disc region because the retinal vessels near the disc region have relatively larger densities than other regions. The examples in Fig. 3 were used to demonstrate the advantages of the proposed density maps generated from the given fundus images.
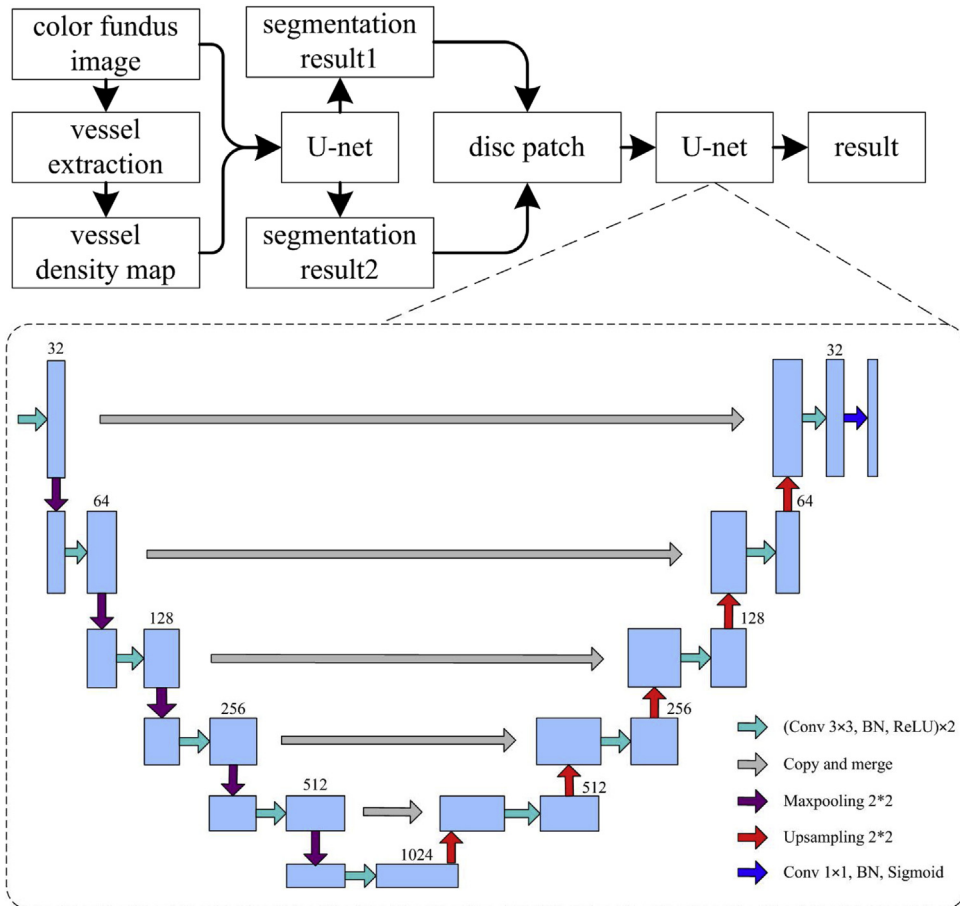
Fig. 2. The flowchart of the proposed method for segmenting the optic disc depicted on color fundus images.

Fig. 3. Illustration of the original fundus images and their vessel density maps. The first column corresponds to the original images, the last two columns to the vessel features and their density maps, respectively.
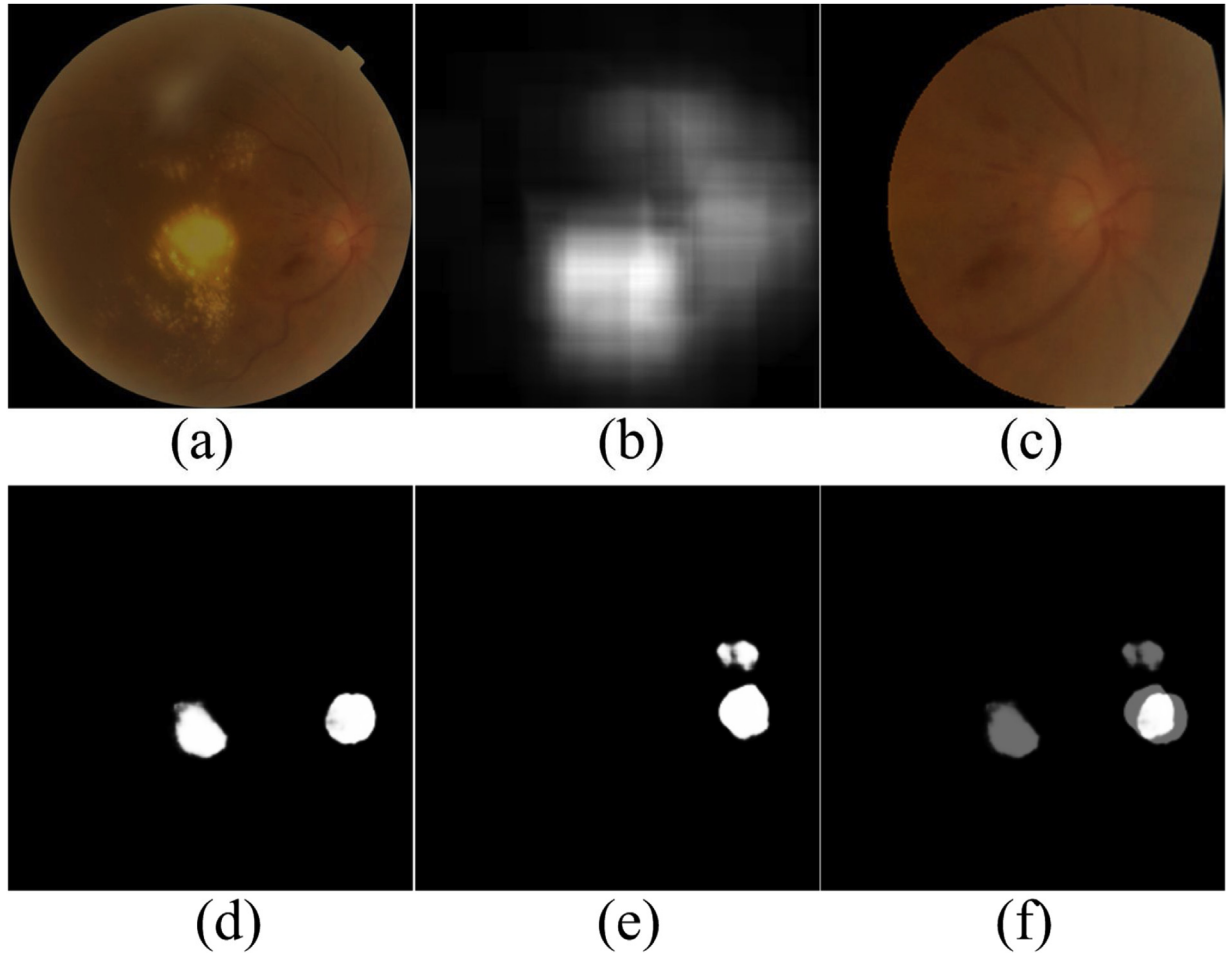
Fig. 4. Example of location of local disc patches by combining two coarse segmentation results. (a) is the original fundus image, (b) is vessel density map of (a), (c) is local disc patch of (a), (d) and (e) are the segmentation results of (a) and (b), respectively, and (f) is the overlapped map of (d) and (e).

It can be seen from the vessel density maps that (1) the vessels around the disc regions typically have high intensity / contrast, (2) the location of the OD is obviously associated with high density values, and (3) existing diseases / abnormalities have very limited impact on the characterization of the OD on the vessel density maps.

2.3. The U-net architecture

With the color fundus images and their vessel density maps, we used a relatively deep U-net convolutional neural network to perform a coarse-to-fine segmentation of the optic disc, as illustrated in Fig. 2. This network learned and captured image features

by stacking two 3×3 convolutional kernels in each convolutional block, followed by a batch normalization (BN)[34] and an element-wise rectified linear unit (ReLU) activation operation [28]. The learned convolution features were down-sampled by leveraging a 2×2 Max-pooling layer with a stride of 2 to reduce the amount of image information. After alternating the convolutional block and Max-pooling layer, the input images were encoded as a large number of high-dimensional features. These features were used to locate the target objects; however, they might fail to precisely specify the boundaries of objects. To obtain both location and boundary information, the high-dimensional features were then decoded by stacking several 2×2 up-sampling layers and convolution blocks mentioned above. During the decoding procedure, a skip connection was employed to add pixel position information and guide the image segmentation. The final decoded results were fed into a 1×1 convolution layer with a sigmoid activation to finalize the target regions.

2.4. Coarse-to-fine segmentation strategy

In implementation, the U-net model was firstly trained on the fundus images with three color channels (i.e., RGB), and their vessel density maps with one grayscale channel, respectively, to generate two different segmentation results. The results were susceptible to relatively large segmentation errors, since the entire images were used and their resolutions were greatly reduced. In addition, the segmentation results often contained false positive regions due to the image quality and / or existing abnormalities. To reduce false positive detection and obtain optimal segmentation results, we combined the above two segmentation results by identifying their overlapped regions and selected the largest one as the most reliable OD candidate region. The selection was based on the fact that most of false positive regions, similar to the OD, were typically caused by some kinds of pathological disease in the original images, and it is very seldom to have these false positive regions on both color fundus images and the vessel density maps. Based on the selection criterion, an image patch centered at the overlapped region was cropped from the original fundus images. The size of the image patch was approximately three times the diameter of an OD to fully enclose the OD and the surrounding regions. Since the OD appears as a circular region, we proposed to use a circular image patch in this study, as shown in Fig. 4. This overlapping strategy can largely exclude the false positive regions, and meanwhile the obtained circular image patch can characterize the OD regions using a relatively high image resolution. Most important, these local image patches improved the contrast between the OD and its surrounding regions as compared to their contrast in the original fundus image. The identified circular image patches with three color channels were then fed into the U-net model again to perform a refined segmentation.

2.5. Training the U-net models

To train the involved models, we collected a total of 7614 color fundus images (CFI) acquired on different subjects in the Shaanxi Provincial People's Hospital, among which there were about 1810 images with weak contrast or severe intensity inhomogeneity. These images have a resolution ranging from 1924×1556 pixels to 2336×2336 pixels. Before the training, we excluded the dark background and resized the field of view from the original fundus images to a dimension of 512×512 pixels. We also normalized the images in order to reduce the impact caused by different illumination conditions, as shown in Fig. 5. These images were manually processed by two experienced ophthalmologists to annotate the disc regions with ImageJ software (https://imagej.net/) using an ellipse. The annotated elliptic regions were used as the ground
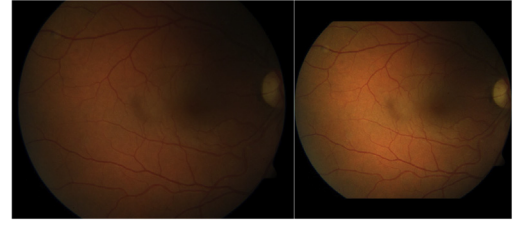

Fig. 5. Example of the original fundus image (left) and its normalized version (right).

truths in this study. These ground truths, together with the normalized fundus images, were randomly divided into three independent parts for training, validation, and testing purposes at a ratio of 0.75, 0.15 and 0.1, respectively, and fed into the above proposed methods for segmenting the optic disc. The three independent datasets contained 1377, 276 and 157 images with weak contrast and severe inhomogeneity, respectively.

The proposed network was implemented using the Keras library (https://keras.io/) and trained on a PC with a 3.50 GHz Intel(R) Xeon(R) E5-1620 CPU, 8GB RAM and NVIDIA Quadro P5000. During the training, we resized all involved images to a dimension of 256×256 pixels and used a batch size of 16 samples, considering the computer capability. We used the intersection over union (IoU) [35] as the loss function, which can be given by:

$$\text{IoU} = \frac{N(A \cap B)}{N(A \cup B)} \qquad (2)$$

where $N(\cdot)$ indicates the number of pixels enclosed set. A and B denote the segmentation results obtained by a specified algorithm and manual annotation, respectively. $\cap$ and $\cup$ denote the intersection and union operators. This function was maximized using the stochastic gradient descent (SGD) with a momentum of 0.95 and an initial learning rate of 0.001. The rate was decreased by a factor of 10 for every 30 epochs. The number of epochs was set at 300 in this study. An early stopping scheme was applied to reduce unnecessary training time if the values of IoU were not improved after 35 epochs. A division point of 0.5 was used to threshold the training results and to decide whether a pixel belongs to the optic disc or not. To prevent overfitting, the fundus images were randomly flipped along each axis, and then rotated [36] from 0 to 360°. These data augmentation techniques increased the diversity of the image samples and were able to significantly improve the generalization and reliability of the deep learning network for these samples.

2.6. Performance evaluation

We validated the developed method on the testing part of our collected CFI images and six publicly available fundus image datasets [1,3,22,33], including DIARETDB0, DIARETDB1, DRIONS-DB, DRIVE, MESSIDOR and ORGIA. These datasets have been widely used for detecting and diagnosing ocular diseases, such as diabetic retinopathy, glaucoma, and age-related macular degeneration. Along with these images, the disc boundaries manually delineated by clinical experts were also available. DIARETDB0 consisted of 130 images, of which 110 had diabetic retinopathy and 20 were normal images. DIARETDB1 contained a total of 89 images, where 84 had mild and non-proliferative diabetic retinopathy. DRIONS-DB was comprised of 110 images, most of which had moderate peripapillary atrophy (PPA) and fuzzy rim. DRIVE contained 40 images, where 33 were normal images and 7 diagnosed with mild diabetic retinopathy. In the MESSIDOR and ORGIA datasets, there were 1200 and 650 fundus images acquired to facilitate the detection of diabetic retinopathy and glaucoma, respectively. In three of these datasets, namely DIARETDB0, DIARETDB1 and MESSIDOR, the images with ocular diseases were annotated. As a result, we

Table 1
Experimental results on the collected CFI images and six public datasets using the U-net model trained separately on the fundus images, vessel density maps, composite images and local disc patches.

Datasets	No. of images	Color fundus images				Vessel density maps			
		IoU	DSC	Acc	Se	IoU	DSC	Acc	Se
CFI	762	0.873	0.923	0.960	0.921	0.687	0.799	0.903	0.809
DIARETDB0	128	0.837	0.898	0.937	0.874	0.684	0.800	0.937	0.878
DIARETDB1	88	0.833	0.904	0.934	0.869	0.686	0.807	0.932	0.869
DRIONS-DB	110	0.818	0.885	0.930	0.862	0.657	0.786	0.859	0.719
DRIVE	40	0.778	0.863	0.895	0.790	0.688	0.802	0.874	0.750
MESSIDOR	1200	0.894	0.940	0.970	0.942	0.739	0.845	0.935	0.874
ORIGA	650	0.864	0.920	0.976	0.954	0.722	0.831	0.938	0.880
Overall	**2978**	**0.874**	**0.925**	**0.964**	**0.930**	**0.714**	**0.824**	**0.924**	**0.851**

Datasets	No. of images	4-channel composite images				Local disc patches			
		IoU	DSC	Acc	Se	IoU	DSC	Acc	Se
CFI	762	0.880	0.928	0.967	0.935	0.886	0.935	0.965	0.935
DIARETDB0	128	0.856	0.916	0.947	0.895	0.848	0.909	0.935	0.872
DIARETDB1	88	0.841	0.905	0.936	0.873	0.849	0.916	0.935	0.873
DRIONS-DB	110	0.870	0.923	0.954	0.909	0.885	0.936	0.960	0.924
DRIVE	40	0.766	0.854	0.889	0.778	0.783	0.864	0.896	0.794
MESSIDOR	1200	0.900	0.943	0.970	0.940	0.907	0.949	0.972	0.948
ORIGA	650	0.873	0.930	0.989	0.979	0.889	0.940	0.989	0.984
Overall	**2978**	**0.882**	**0.931**	**0.969**	**0.939**	**0.891**	**0.939**	**0.970**	**0.944**

Table 2
Experimental results on the normal group and the abnormal group from DIARETDB0, DIARETDB1, and MESSIDOR using the U-net model trained separately on the fundus images, vessel density maps, composite images and local disc patches.

Datasets	No. of images	Color fundus images				Vessel density maps			
		IoU	DSC	Acc	Se	IoU	DSC	Acc	Se
Normal	571	0.875	0.927	0.959	0.920	0.740	0.845	0.936	0.876
Abnormal	845	0.896	0.942	0.970	0.940	0.725	0.834	0.935	0.873
Overall	**1416**	**0.885**	**0.934**	**0.965**	**0.931**	**0.731**	**0.838**	**0.935**	**0.874**

Datasets	No. of images	4-channel composite images				Local disc patches			
		IoU	DSC	Acc	Se	IoU	DSC	Acc	Se
Normal	571	0.899	0.943	0.968	0.938	0.904	0.947	0.970	0.944
Abnormal	845	0.887	0.935	0.964	0.928	0.894	0.941	0.964	0.932
Overall	**1416**	**0.892**	**0.939**	**0.966**	**0.932**	**0.898**	**0.944**	**0.967**	**0.937**

setup a normal group consisting of 845 images and an abnormal group consisting of 571 images to evaluate the performance of the developed approach in the presence of disease.

With these publicly available datasets, we compared the developed method with available OD segmentation methods using different performance metrics. These metrics included the IoU, the dice similarity coefficient (DSC), and balanced accuracy (Acc) [3], which are defined, respectively, as:

$$DSC = \frac{2N(A \cap B)}{N(A) + N(B)} \quad (3)$$

$$Acc = \frac{Se + Sp}{2} \quad (4)$$

$$Se = \frac{Tp}{Tp + Fn} \quad (5)$$

$$Sp = \frac{Tn}{Tn + Fp} \quad (6)$$

where Tp, Tn, Fp and Fn are the number of true positives and true negatives, false positives and false negatives, respectively. Se and Sp [18] denote the sensitivity and specificity of a given segmentation method, the former is used to measure the desired pixels with positive values while the latter measures the undesired pixels with negative values. These metrics take values in the range from 0 to 1, and a higher value means better segmentation accuracy. In addition, we used the disc location accuracy (DLA) and the mean

distance to closest point (MDCP) [7] for performance comparison. The DLA was defined as the ratio of the number of correctly estimated disc centers and the total number of tested images. Similar to [25], we considered the estimated OD center correct if it was within 60 pixels of the manually identified center. MDCP was used to characterize the degree of mismatch between the segmented contour and the ground truth contour. Analysis of variance (ANOVA) was performed to assess the statistical differences among these involved methods, where we considered a p-value less than 0.05 statistically significant.

3. Results

Tables 1 and 2 show the segmentation results of the collected CFI images and six public datasets using the U-net model trained on color fundus images, vessel density maps, 4-channel composite images (comprised of the original RGB image and its vessel density map), and the local disc image patches, respectively. The values of IoU, DSC, Acc and Se obtained by our method in Table 1 were, on average, higher by 1.73%, 1.40%, 0.59% and 1.51%, respectively, than the counterparts of the U-net model trained on the same set of original fundus images. The p-values for these four metrics were less than 0.001. Our method also had larger average values of IoU and DSC than those of the U-net model trained on 4-channel composite images along with the p-value of 0.001. This suggested that there were statistically significant differences between the pro-
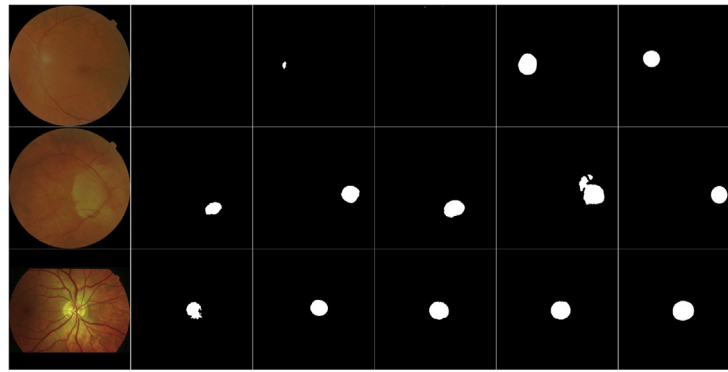
Fig. 6. Illustration of segmentation results of the U-net model on different images. The first column showed three original fundus images, the last five columns from left to right were results of these fundus images, vessel density maps, 4-channel composite images, local disc patches, and their ground truths, respectively.

posed method and the U-net method trained separately on the original image and 4-channel composite images. The sole U-net model had the lowest segmentation accuracy when trained on the grayscale vessel density maps, as compared to the color fundus images, suggesting that the developed method was superior to the U-net model trained merely on the original fundus images or their vessel density maps. In particular, our method in Table 2 achieved the average IOU and DSC of 0.894 and 0.941 for the abnormal group, and 0.904 and 0.947 for the normal group, respectively. In contrast, the counterparts of the U-net model were 0.875 and 0.927 for the abnormal group, 0.896 and 0.942 for the normal group.

Some examples in Fig. 6 demonstrate the segmentation results by these methods. The OD regions were not effectively identified from the original images using the U-net model when these images had severe intensity inhomogeneity and / or abnormalities, while the OD regions were successfully located from the local disc patches. The proposed method took an average of 0.04 s to segment the OD for a single image.

We also trained the U-net model on 4-channel disc patches (comprised of local disc regions and their local vessel density maps) in the fine segmentation step, and achieved the average IoU, DSC and Acc of 0.890, 0.939, and 0.966, respectively, on the same testing datasets. The results were slightly inferior to those (i.e., 0.891, 0.939 and 0.970) of our method trained on local disc patches with RGB channel, but there were no significant differences between the two results.

Table 3 summarizes a performance comparison of the proposed method with several available OD segmentation methods [37–42] based on six public datasets. As demonstrated by the results, the proposed method achieved, on average, 99.82% disc location accuracy and 2.47 pixel distance to closest contour point in ground truths, demonstrating a better performance than available approaches.

4. Discussion

We developed a coarse-to-fine deep learning framework on the basis of the U-net model to segment the OD depicted on color fundus images. The most unique characteristic of the proposed method was the integration of two types of image information (i.e., pixel intensities and vessel density) to locate and segment the OD. In particular, we developed a vessel density map to characterize the spatial relationship between the OD and retinal vessels, to help detect the OD and to exclude many false positive objects in the entire image region. This density map was especially useful for the images where the contrast between the OD and the surrounding region was not good but the vessels were well depicted on the images.

Table 3
Performance comparison of different disc segmentation methods on six public datasets using the metrics of IoU, Se, DLA, and MDCP (pixels), respectively. The dash ("-") denotes unavailable measures in related literature.

	IoU	Se	DLA	MDCP
DIARETDB0				
Kumar et al [7]	89.6%	93.3%	–	8.62
Zhang et al [5]	–	–	95.5%	–
Sinha et al [10]	–	–	96.9%	–
Our method	84.8%	87.2%	98.4%	4.97
DIARETDB1				
Welfer et al [8]	43.7%	–	97.8%	8.31
Abdullah et al [22]	85.1%	85.1%	100.0%	–
Walter et al [13]	37.3%	–	–	15.52
Seo et al [38]	35.3%	–	–	9.74
Lupascu et al [39]	31.0%	–	–	13.81
Kande et al [14]	33.2%	–	–	8.35
Morales et al [16]	81.7%	92.2%	99.6%	2.88
Our method	84.9%	87.3%	100.0%	2.75
DRIONS-DB				
Morales et al [16]	84.2%	92.8%	99.34%	2.49
Maninis et al [27]	88.0%	–	–	–
Kumar et al [17]	89.0%	–	–	–
Abdullah et al [22]	85.1%	85.1%	99.09%	–
Our method	88.5%	92.4%	100.0%	2.59
DRIVE				
Walter et al [13]	29.3%	49.9%	–	12.39
Seo et al [38]	31.1%	50.3%	–	11.19
Abdullah et al [22]	78.6%	81.9%	100.0%	–
Morales et al [16]	71.6%	85.4%	99.0%	5.85
Lalonde et al [40]	80.0%	–	–	–
Our method	78.3%	79.4%	97.5%	9.57
MESSIDOR				
Aquino et al [9]	86.0%	–	99.0%	–
Lu et al [36]	–	–	99.8%	–
Kumar et al [7]	91.6%	–	–	–
Abdullah et al [22]	87.9%	89.5%	99.3%	–
Morales et al [16]	82.3%	93.0%	99.5%	4.08
Our method	90.7%	94.8%	99.9%	2.10
ORIGA				
Yin et al [4]	88.7%	–	–	–
Cheng et al [41]	90.0%	–	–	–
Cheng et al [42]	90.5%	–	–	–
Fu et al [3]	92.9%	–	–	–
Our method	88.9%	98.4%	100.0%	2.17

We quantitatively assessed the segmentation performance of the proposed method using a total of 2978 images from the collected CFI dataset and six public datasets. The extensive validation on multiple datasets demonstrates the effectiveness and robustness of the proposed method against different fundus cameras, imaging conditions, image resolutions, and abnormalities. In contrast, the existing methods were typically assessed using a small
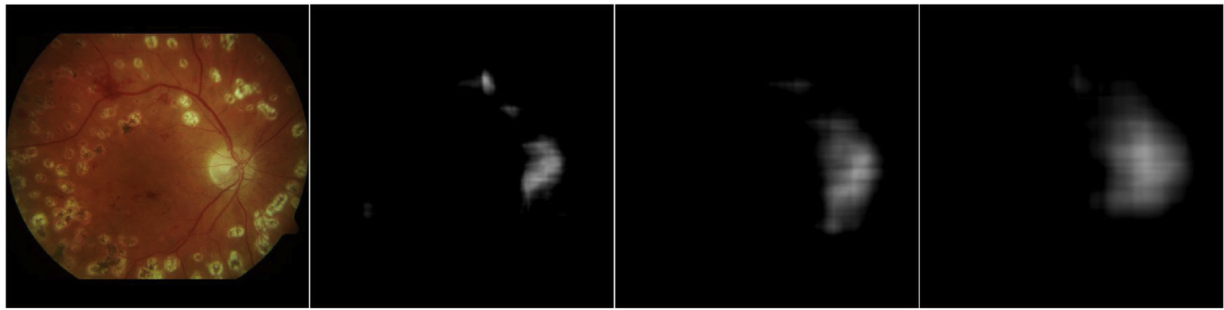
Fig. 7. Illustration of the original image and its vessel density maps calculated with three different *R* values of 20, 40 and 60 pixels, respectively.

number of test images. For example, Kumar et al. [17] only used 112 images, of 130 images, from the DIARETDB0 dataset and 40 images, of 1200 images, from the MESSIDOR dataset. Yin et al. [4] only tested 325 images, of 650 images, from the ORIGA dataset. When using the same testing images as those in paper [4], the proposed method achieved the IoU and DSC of 0.885 and 0.938, respectively, which were similar to those of the method [4] (i.e., 0.887 and 0.941). In addition, tests on the normal and abnormal groups suggested that the developed method was not very sensitive to the presence of ocular diseases (Table 2). It is notable that due to the hardware (GPU memory) limitation, our method was trained on images with a dimension of 256×256 pixels. In contrast, most of available deep learning methods [3,27,28] were trained on images with relatively large dimensions. For example, the images were resized to 400×400 pixels and 565×584 pixels, respectively, for training purpose in [3] and [27]. Hence, given the size and diversity of our testing dataset, our method demonstrated a relatively reliable and reasonable performance.

The proposed method involved several important parameters, among which *R* and the local patch size were important. *R* determines the size of the disc region as depicted on the density map, while the local image patch size may affect segmentation performance. In this study, the *R* value was approximately equal to the radius of the optic disc. When R was too small, the vessel density map only highlighted the vessels and only a small part of the optic disc, as demonstrated by the example in Fig. 7, where three different vessel density maps were calculated using the *R* values of 20, 40 and 60 pixels, respectively. When R was too large, the OD boundaries were not well characterized on the density map. The local patch size was set at a value approximately equal to three times the diameter of the optic disc, aiming to fully enclose the OD regions and provide a relatively high resolution of the OD region, and meanwhile alleviate a high requirement on computer hardware (e.g., the GPU memory).

Finally, there are some limitations with this study. First, the ground truth for the optic disc may not be perfectly delineated in color fundus images due to the influence of the peripapillary atrophy [1] and subjective bias. This makes the U-net model potentially unable to locate desirable object boundaries when trained on local disc patches, since the peripapillary atrophy and fuzzy rim were enlarged and might affect the estimation of the pixels within the disc regions, leading to reduced segmentation accuracy. Second, the accuracy of the blood vessel identification may affect the results as well. When the vessels are not well depicted on the images or there are severe diseases, the vessel density maps may fail to highlight the position of the optic disc, and introduce irrelevant object regions in image segmentation. Third, we selected a low image resolution (i.e., 256×256) in this study with consideration of the limitation of the GPU hardware in memory. This resolution might limit the amount of image features in the deep learning network. Fu et al [3] and Maninis et al [27] trained different variants of the U-net model on images with a resolution of 400×400 and 584×565, respec-

tively, which were much larger than the counterpart used in our experiment. Despite these limitations, the proposed method can automatically identify the OD regions with very promising accuracy.

5. Conclusion

In this paper, we introduced a coarse-to-fine deep learning framework based on the U-net model to accurately and automatically segment the OD depicted on color fundus images. Its novelty mainly lies in integrating different image information (i.e., image contrast and retinal vessels) and leveraging a coarse-to-fine segmentation strategy. Our extensive experiments on the collected images and six public datasets, which consisted of a total of 2978 fundus images, demonstrated that the proposed framework could achieve a relatively reliable segmentation performance.

Acknowledgements

References

[1] A. Almazroa, R. Burman, K. Raahemifar, V. Lakshminarayanan, Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey, J. Ophthalmol. 2015 (2015) 1–28.

[2] A. Singh, M. Dutta, M. Sarathi, V. Uher, R. Burget, Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image, Comput. Methods Programs Biomed. 124 (2016) 108–120.

[3] H. Fu, J. Cheng, Y. Xu, D. Wong, J. Liu, X. Cao, Joint optic disc and cup segmentation based on multi-label deep network and polar transformation, IEEE Trans. Med. Imaging (2018), http://dx.doi.org/10.1109/TMI.2018.2791488.

[4] F. Yin, J. Liu, S. Ong, et al., Model-based optic nerve head segmentation on retinal fundus images, in: Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2011, pp. 2626–2629.

[5] D. Zhang, Y. Yi, X. Shang, Y. Peng, Optic disc localization by projection with vessel distribution and appearance characteristics, in: International Conference on Pattern Recognition (ICPR), 2012, pp. 3176–3179.

[6] J. Zilly, J. Buhmann, D. Mahapatra, Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation, Comput. Med. Imaging Graph. 55 (2017) 28–41.

[7] V. Kumar, N. Sinha, Automatic optic disc segmentation using maximum intensity variation, in: Proceedings of the IEEE TENCON Spring Conference, 2013, pp. 29–33.

[8] D. Welfer, J. Scharcanski, C. Kitamura, M. Dal Pizzol, L. Ludwig, D. Marinho, Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach, Comput. Biol. Med. 40 (2) (2010) 124–137.

[9] M. Aquino, E. Gegundez, D. Mar, Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques, IEEE Trans. Med. Imaging 29 (11) (2010) 1860–1869.

[10] N. Sinha, R. Babu, Optic disk localization using L1 minimization, in: Proceedings of the 19th IEEE International Conference on Image Processing (ICIP), 2012, pp. 2829–2832.

[11] L. Wang, Y. Chang, H. Wang, Z. Wu, J. Pu, X. Yang, An active contour model based on local fitted images for image segmentation, Inf. Sci. 418 (2017) 61–73.

[12] L. Wang, G. Chen, D. Shi, Y. Chang, S. Chan, J. Pu, X. Yang, Active contours driven by edge entropy fitting energy for image segmentation, Signal Processing 149 (2018) 27–35.

[13] T. Walter, J. Klein, P. Massin, A. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina, Trans. Med. Imaging 21 (10) (2002) 1236–1243.

[14] G. Kande, P. Subbaiah, T. Savithri, Segmentation of exudates and optic disc in retinal images, in: Sixth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), 2008, pp. 535–542.

[15] H. Alghamdi, H. Tang, S. Waheeb, T. Peto, Automatic optic disc abnormality detection in fundus images: a deep learning approach, in: Proceedings of the Ophthalmic Medical Image Analysis (OMIA), 2016, pp. 17–24.

[16] S. Morales, V. Naranjo, J. Angulo, M. Alcaniz, Automatic detection of optic disc based on PCA and mathematical morphology, IEEE Trans. Med. Imaging 32 (2013) 786–796.

[17] J. Kumar, A.Pediredla C. Seelamantula, Active discs for automated optic disc segmentation, in: 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2015, pp. 225–229.

[18] L. Wang, H. Zhang, K. He, Y. Chang, X. Yang, Active contours driven by multi-feature Gaussian distribution fitting energy with application to vessel segmentation, PLoS One 16 (2015).

[19] A. Giachetti, L. Ballerini, Trucco E, Accurate and reliable segmentation of the optic disc in digital fundus images, J. Med. Imaging 1 (2) (2014), 024001.

[20] M. Hu, C. Zhu, X. Li, Y. Xu, Optic cup segmentation from fundus images for glaucoma diagnosis, Bioengineered 8 (1) (2017) 21–28.

[21] A. Gopalakrishnan, A. Almazroa, K. Raahemifar, L. Lakshminarayanan, Optic disc segmentation using circular hough transform and curve fitting, in: Conference on Opto-Electronics Applied Optics, 2015, pp. 1–4.

[22] M. Abdullah, M. Fraz, S. Barman, Localization and segmentation of optic disc in retinal images using circular Hough transform and grow-cut algorithm, PeerJ (2016).

[23] J. Sigut, O. Nunez, F. Fumero, M. Gonzalez, R. Arnay, Contrast based circular approximation for accurate and robust optic disc segmentation in retinal images, PeerJ (2017).

[24] A. Hoover, M. Goldbaum, Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels, IEEE Trans. Med. Imaging 22 (8) (2003) 951–958.

[25] A. Youssif, A. Ghalwash, A. Ghoneim, Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter, IEEE Trans. Med. Imaging 27 (1) (2008) 11–18.

[26] J. Tan, U. Acharya, S. Bhandary, K. Chua, S. Sivaprasad, Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network, J. Comp. Sci. 20 (2017) 70–79.

[27] K. Maninis, J. Tuset, P. Arbelaez, L. Gool, Deep retinal image understanding, in: Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2016, pp. 140–148.

[28] G. Litjens, T. Kooi, B. Bejnordi, A. Setio, F. Ciompi, M. Ghafoorian, J. Laak, B. Ginneken, C. Sanchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–68.

[29] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv:1505.04597v1, 2015.

[30] A. Budai, J. Hornegger, G. Michelson, Blood vessel density and distribution in fundus images, Invest. Ophthalmol. Vis. Sci. 53 (14) (2012) 2175.

[31] F. Milletari, N. Navab, S. Ahmadi, V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, arXiv:1606.04797v1, 2016.

[32] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.

[33] Y. Zhen, S. Gu, X. Meng, X. Zhang, N. Wang, J. Pu, Automated identification of retinal vessels using a multiscale directional contrast quantification (MDCQ) strategy, Med. Phys. 41 (9) (2014), 092702.

[34] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, arXiv preprint arXiv:1502.03167v3, 2015.

[35] L. Wang, J. Zhu, M. Sheng, A. Cribb, S. Zhu, J. Pu, Simultaneous segmentation and bias field estimation using local fitted images, Pattern Recognit. 74 (2018) 145–155.

[36] L. Wang, X. Gao, X. Cui, Z. Liang, 2D/3D rigid registration by integrating intensity distance, Opt. Precis. Eng. 22 (10) (2014) 2815–2824.

[37] S. Lu, Accurate and efficient optic disc detection and segmentation by a circular transformation, IEEE Trans. Med. Imaging 30 (12) (2011) 2126–2133.

[38] J. Seo, K. Kim, J. Kim, K. Park, H. Chung, Measurement of ocular torsion using digital fundus image, in: Proceedings of the 26th Annual International Conference of the IEEE EMBS, 2004, pp. 1711–1713.

[39] C. Lupascu, D. Tegolo, L. Rosa, Automated detection of optic disc location in retinal images, in: Proceedings of the 21st IEEE International Symposium on Computer-Based Medical Systems (CBMS), 2008, pp. 17–22.

[40] M. Lalonde, M. Beaulieu, L. Gagnon, Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching, IEEE Transactions on Medical Imaging 20 (11) (2001) 1193–1200.

[41] J. Cheng, J. Liu, D. Wong, et al., Automatic optic disc segmentation with peripapillary atrophy elimination, in: Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2011, pp. 6224–6227.

[42] J. Cheng, J. Liu, Y. Xu, et al., Superpixel classification based optic disc and optic cup segmentation for glaucoma screening, IEEE Trans. Med. Imaging 32 (6) (2013) 1019–1032.