Oval Shape Constraint based Optic Disc and Cup Segmentation in Fundus Photographs

Jun Wu[1]
junwu@nwpu.edu.cn
Kaiwei Wang[1]
wangkaiwei@mail.nwpu.edu.cn
Zongjiang Shang[1]
shangzongjiang@mail.nwpu.edu.cn
Jie Xu (Corresponding Author)[2]
fionahsu920@foxmail.com
Dayong Ding[3]
dayong.ding@vistel.cn
Xirong Li[4]
xirong@ruc.edu.cn
Gang Yang[5]
yanggang@ruc.edu.cn

[1] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China

[2] Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Science Key Lab, Beijing, China

[3] Vistel AI Lab, Visionary Intelligence Ltd., Room 1703, Cultural Building, Zhongguancun Avenue No.59, Haidian District, Beijing, China

[4] School of Information Science and Technology, University of Science and Technology of China, Hefei, China

[5] School of Information, Renmin University of China, Beijing, China

Abstract

In a fundus photograph, morphological changes of the optic disc and cup are crucial for diagnosing optic neuropathy. To achieve an accurate pixel-wise segmentation of the optic disc and cup, the domain-specific knowledge such as the oval shape constraint has not been sufficiently explored in most of the existing methods, leading to unacceptable geometric distortions in many cases. Few attempts try to consider the general convexity constraint or specific building geometric properties, but they are still not suitable for the typical oval shape segmentation. In this paper, an oval shape constraint based loss function (OS-loss) is proposed to improve the existing deep learning network for segmenting optic disc and cup. A penalty point set is proposed to represent unreasonable contour points of a target object using the oval shape constraint. These points will be penalized and integrated into the training loss function of the baseline network. Further, an oval-friendly metric called shape error (SE) is proposed to better reflect the fitness of two oval contours. Experiments on the public RIM-ONE-r3 dataset with 159 fundus photographs and a private W10K dataset with 9,879 fundus photographs prove the effectiveness of the proposed OS-loss function. Compared to the original CE-net, the mean error of the Cup to Disc Ratio (CDR) of the proposed OS-loss method in the RIM-ONE-r3 dataset decreases 1.98%. In the W10K dataset, the mean CDR error decreases by 1.03% for the ResU-net and decreases by 2.1% for the CE-net.

1 Introduction

The automatic segmentation of the optic disc and cup is crucial for further diagnosing retinal optic nerve diseases. The shape of the optic disc (OD) and optic cup (OC), their morphological changes, as well as the Cup to Disc Ratio (CDR), are playing an important role in the diagnosis of ophthalmic diseases, e.g., glaucoma [1]. A glaucoma is a group of eye diseases that damage the human eye optic nerve and it can result in vision loss and blindness irreversibly. Automatic CDR estimation requires an accurate segmentation of optic disc and cup. According to the anatomical structure of the human eyes, the optic disc and cup are oval-shaped, and an optic cup is usually contained within an optic disc. With these constraints, the oval-shaped prior knowledge should be sufficiently explored. Besides, oval shapes are popular in medical images, such as cells, cataract nuclear regions, pupils and iris.

To the best of our knowledge, there are three branches of the optic disc and cup segmentation methods: (1) Image processing [8] applies the RGB color information, threshold method [11], or C-V model [3] based segmentation, etc. (2) Traditional machine learning utilizes classical classifiers based on typical image features, such as support vector machine [2], neural network [18], etc. (3) Deep learning is widely employed to segment optic disc and cup. The U-net [14] uses end-to-end training for pixel-wise image segmentation. On the small training set, it has satisfactory performance and high training speed. The ResU-net [16] combines ResNet[7] blocks with the U-net framework. It uses residual error and block coding and decoding operations to make a deeper network structure. The CE-net[6] proposed by Gu *et al*. combines the feature encoder (the pre-trained ResNet block), context extractor (the dense atrous convolution and residual multi-kernel pooling block) and feature decoder to capture high-level and spatial information for medical image segmentation tasks.

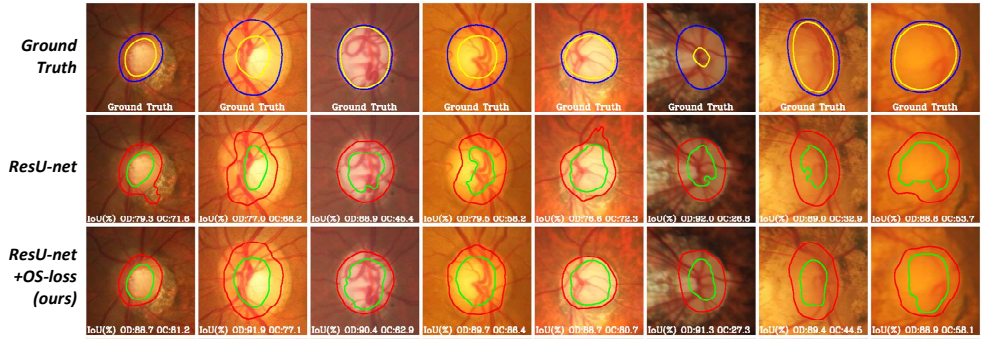

Figure 1: Failure ROI examples of segmenting optic disc and cup by the ResU-net [16] and our proposed oval shape constraint based loss function (ResU-net+OS-loss). For the contour results, optic disc: in the red, optic cup: in the green. For the ground truth, optic disc: in the blue, optic cup: in the yellow. The overlap rate IoU(%) is defined in Section 3.2.

For the pixel-wise image segmentation, due to the restrictions of the perception field of the existing deep learning networks such as a fully convolutional network (FCN) [9], U-net [14] or ResU-net [16] , in many cases the global shape information of an object cannot be sufficiently explored, leading to unacceptable geometric distortions and poor segmentation results, as shown in Figure 1. Though there are some existing works consider the convexity constraint [6, 15, 17] or specific building geometric properties [10] in an image segmentation task, they are still not suitable for the oval shape segmentation such as optic disc and

cup. Ravishankar *et al*. incorporate shape priors within the FCN segmentation framework [12], however, its shape regularization network requires additional training data with shape annotations in expensive costs, and it is either not tailored for oval-shape segmentation.

In this paper, an oval shape constraint based loss function is proposed to better segment optic disc and cup in color fundus photographs. It sufficiently explores the domain knowledge such as oval shape constraint to guide the network training and optimization to obtain more reasonable and stable contours of the optic disc and cup.

2 OS-loss: The Proposed Penalty Loss Function

We design a novel loss function based on the oval shape constraint to segment the optic disc and cup. As illustrated in Figure 2, there are three main modules in our framework: the optic disc ROI localization, the baseline deep learning network, and the oval shape constraint based loss function (hereinafter, OS-loss) updating. First, the optic disc ROI is automatically localized from the original fundus photograph using the Faster R-CNN [13]. Then, popular deep learning networks such as the FCN [9], U-net [14] , ResU-net [16] or CE-net [6] can be applied. Finally, the OS-loss penalty function will be implemented base on the shape constraint module, which will update the current training loss function.

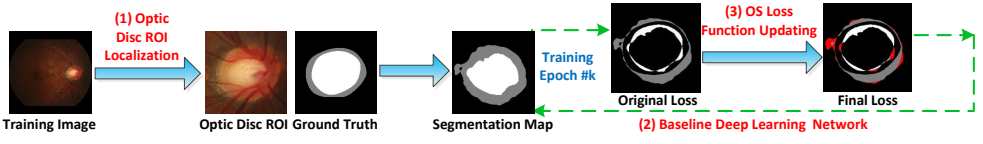

Figure 2: An overview flowchart of our proposed OS-loss function in a baseline network.

As illustrated in Figure 3, the OS-loss function updating module includes four main steps: the contour point validation, the contour point optimization, the penalty point set verification and the loss weight matrix updating. At a certain training epoch, the contour points of the optic disc and cup from a segmentation mask are first verified based on the global contour oval shape constraint. Then, an ellipse fitting is applied base on these remaining valid contour points to achieve an optimized contour point set. Next, a penalty point set with unreasonable contour points is formed based on the above-fitted ellipse boundary. Finally, these penalty points are assigned as proper weights to update the training loss function.

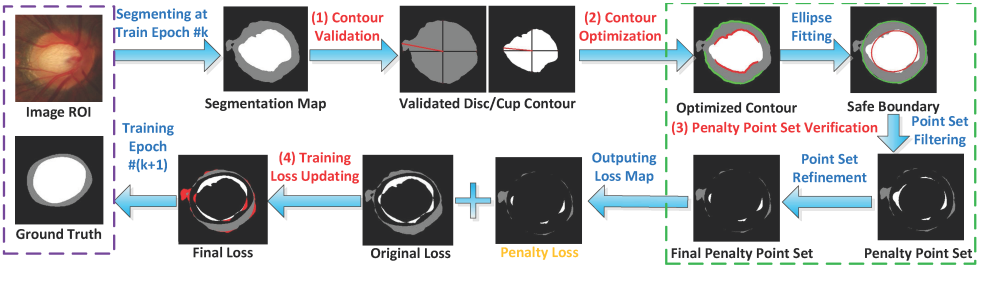

Figure 3: A detailed flowchart of our proposed OS-loss function module.

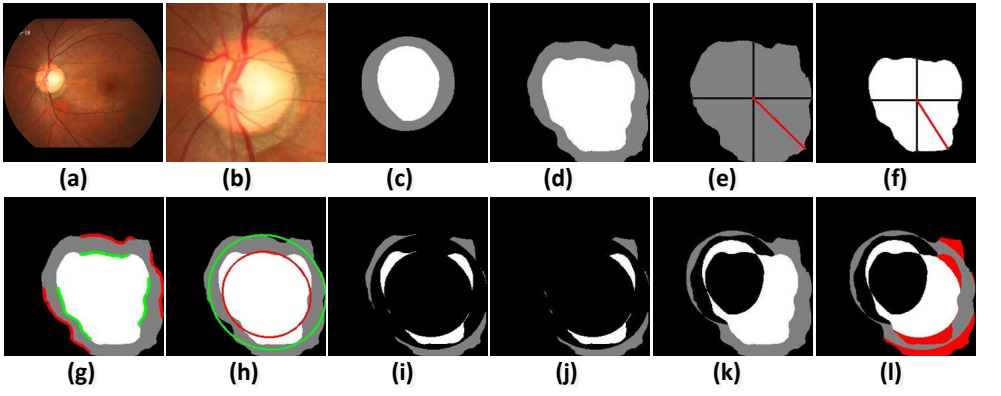

Figure 4: An example of the workflow of the shape constraint module for optic disc and cup: (a) the original fundus photograph, (b) the optic disc ROI, (c) the ground truth mask, (d) the segmentation results in a certain training epoch, (e) optic disc contour point validation, (f) optic cup contour point validation, (g) the optimized contour, (h) the ellipse fitting results, (i) the penalty point set, (j) the verified penalty point set, (k) original loss weight matrix, (l) final loss weight matrix, considering penalty points in a red color.

2.1 The Contour Point Validation

We use an optic cup as an example. Supposing the ROI of a given training image I is in $M \times N$ pixels, as in Figure 4(b). The contour of the optic cup is obtained from its corresponding segmentation map S at a certain batch in the k-th training epoch, as in Figure 4(d).

Then, for the region point set of the optic cup R surrounded by the predicted contour point set $C \subset R$, its geometric center of gravity is calculated as the coordinate origin $O=(0,0)$. Four intersection points between the object contour and both the x-axis and y-axis are: x_1, x_2, y_1 and y_2. To roughly reflect the average oval shape size of the optic cup in both horizontal and vertical directions, an average shape distance $\bar{D}$ is defined as:

$$\bar{D} = \frac{1}{4}(|x_1| + |x_2| + |y_1| + |y_2|). \tag{1}$$

Further, if a contour point is valid, it means it is reasonable based on the oval shape constraint. As in Figure 4(e)(f), the validity of the target contour point $V(p)$ is defined as:

$$V(p) = \begin{cases} 1, & p \leq \bar{D} + \varepsilon, \\ 0, & otherwise, \end{cases} \tag{2}$$

where p is a random point in the contour point set C and ε is a controlling threshold introducing a reasonable shape distortion.

2.2 The Contour Point Optimization

A strategy of dynamic tuning of the threshold ε as in Section 2.1 is applied to obtain an optimized point set as in Figure 4(g), which is between a reasonable percentage of the original contour point set, *i.e.,* between 50% and 70%. At the beginning ε is set as 0. According to Equation (2), some outliers are removed and a new contour point set is formed. If the current

portion of the remaining points is smaller than 50% of the original point set, set $\varepsilon = 0.01\bar{D}$. If it is bigger than 70%, then set $\varepsilon = -0.01\bar{D}$. The same process will be repeated until the size of the new point set is between the desired portion of the original contour point set.

2.3 The Penalty Point Set Verification

The ellipse fitting is carried out on the optimized contour point set to generate an ellipse as in Figure 4(h), which is employed as an inside-safe boundary to further determine a penalty point set, as in Figure 4(i), where all the points are unreasonable. All the optic cup region points inside this ellipse are considered as a valid point set R_1, then all the other optic cup region points outside it are filtered as the penalty point set R_2, which will be used to update the original training loss weight matrix as in Figure 4(k). Finally, the correctly-predicted pixels (i.e., the prediction is the same as the ground truth) will be removed from the penalty point set R_2, resulting in a verified point set $R_2^* = \{p_i | l_i(p_i) \neq \bar{l}_i(p_i), p_i \in R_2\}$ as in Figure 4(j), where $\bar{l}_i$ is the class label of ground truth for the pixel p_i and l_i is its predicted class label.

2.4 The Training Loss Weight Matrix Updating

As in Figure 4(l), the verified penalty point set R_2^* can be utilized to update the original training loss function to penalize possible similar mistakes, e.g., the bigger weights will be applied. Thus, the newly-updated loss function L^* can be determined as:

$$L^* = W_{M\times N} \cdot L^{CE}_{M\times N} = \sum_{i=1}^{M\times N} (w_i \times loss_i^{CE}), w_i \in W_{M\times N} \text{ and } loss_i^{CE} \in L^{CE}_{M\times N}, \tag{3}$$

where the updated weight of w_i for the loss function is defined as:

$$w_i = \begin{cases} 1 + w^*, & p_i \in R_2^* (\text{verified penalty point set}), \\ 1, & otherwise. \end{cases} \tag{4}$$

Note: w^* is a penalty weight. The pixel-wise $loss_i^{CE}$ is defined as $loss_i^{CE} = -\bar{l}_i \log(l_i)$.

3 Evaluation

3.1 Experimental Setup

We validate our proposed OS-loss function on two datasets: (1) RIM-ONE-r3 [9]: a public dataset including 159 stereo retinal fundus photographs. Only the monocular image from the left part is used here. There are 111 images used for training, 24 images for validation, and 24 images for testing. (2) W10K: a private dataset with 9,879 fundus photographs collected from our local hospital partner. The boundaries of optic disc and cup are manually labeled by the professional ophthalmologists. As in Table 1, it is partitioned into two non-overlapping datasets, S1 and S2, which are further divided into the training, validation and testing sets of independent patients. As S1 is small and the potential samples with obvious shape distortion are very few, it is not stable for the parameter optimization. For S1 and S2, we apply the same validation and testing sets in the entire W10K dataset.

Experiments are conducted with an NVIDIA 1080Ti GPU. The disc ROIs are all resized into 256×256. The batch size is 8, the learning rate is $1/10^6$, and $K = 360$.

Table 1: An overview of our private data set W10K. S1 and S2 are two non-overlapping datasets, and they are also non-overlapping for the training, validation and testing subsets.

Data Set	Training(70%)	Validation(15%)	Testing(15%)	Total
S1	**525**	113	114	752
S2	**6,391**	1,368	1,368	9,127
W10K(Total)	6,916	**1,481**	**1,482**	9,879

3.2 Evaluation Metrics

Typical image segmentation metrics such as IoU, $F1$ and Dice coefficient are applied. IoU are the overlap rate, *i.e.,* the coincidence degree between predicted results and the ground truth, as $IoU = |S_o \bigcap S_G| / |S_o \bigcup S_G|$, where S_o is the predicted region, and S_G is the ground truth. $F1 = 2 \times \text{precison} \times \text{recall} / (\text{precison} + \text{recall})$. $Dice = 2|S_o \bigcap S_G| / (|S_o| + |S_G|)$. Specially, CDR is the optic cup to disc ratio based on vertical distances. E_{CDR} is defined as the absolute difference between predicted CDR and ground truth CDR_g, i.e., $|CDR - CDR_g|$.

However, the widely-used metrics IoU mainly computes the overlapping ratio between the segmentation results and the ground truth, and F1 is a metric from information retrieval. Neither of them can reflect how good the predicted shape is fitted into the ground truth. Even if a predicted contour has an unreasonable shape, it also might have a high IoU or F1 score.

Thus, we propose a new metric for the oval shape segmentation, called Shape Error (SE). It mainly reflects the accumulated differences between two oval shapes, *i.e.,* the segmentation result and the ground truth. Similar to the idea of computing $\bar{D}$ in Equation (1), the contour geometric center is O and the object contour point set is C. Supposing C' is the contour point set of the ground truth. A ray $\lambda_{i,\theta}$ ($i = \{1, 2, \cdots, K\}$, and $\theta = i \times 360/K$ is the degree in the polar coordinate system) from O has two intersection points, $c_i = (x, y)$ and $c'_i = (x', y')$, with the result contour and the ground truth contour, as illustrated in Figure 5(f).

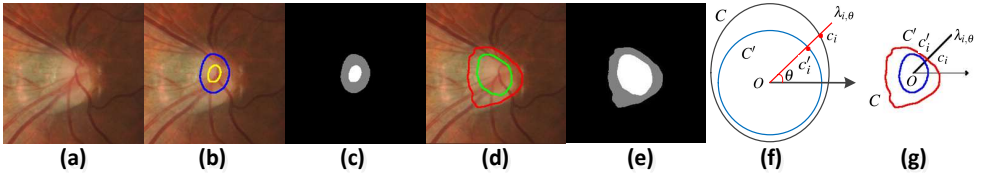

Figure 5: Main ideas of an oval-friendly segmentation metric. (a) an optic disc ROI example, (b) the ground truth contour of the optic disc and cup, (c) the ground truth mask, (d) the contour of the optic disc and cup, (e) the segmentation map, (f) a sketch example of $SE(i)$ in the polar system, (g) the optic disc example of the shape error $SE(i)$ at the line $\lambda_{i,\theta}$.

Finally, in order to use an accumulative normalized error to reflect the fitness degree between oval-shaped contours, the oval shape error (SE) is defined as: $SE = \frac{1}{K} \sum_{i=1}^{K} SE(i) / \bar{D}$, where the specific shape error distance $SE(i)$ at the line $\lambda_{i,\theta}$ is defined as: $SE(i) = D(c_i, c'_i) = D[(x, y), (x', y')] = \sqrt{(x - x')^2 + (y - y')^2}$.

3.3 Resutls

3.3.1 Optimizing the best adding time of the shape constraint module

We first discuss the best time to add our oval shape constraint module into the baseline network. It is added at different training epochs, as listed in Table 2. We choose the best adding time as the 150-th training epoch, as this setting achieves the best performance in terms of most metrics, and only the metrics for optic disc achieves the second best performance with little performance loss (IoU: 0.04%, F1: 0.08%, and Dice: 0.03%).

Table 2: The parameter optimization for the adding time of the ResU-net plus OS-loss. The training set is S1-training (525 photographs) and the validation set is W10K-validation.

No.	Add #iter	Optic Disc				Optic Cup				E_{CDR}
		IoU	F1	Dice	SE	IoU	F1	Dice	SE	
1	30	86.89	93.16	92.71	5.56	72.06	85.29	82.64	16.59	9.95
2	50	**88.24**	**94.09**	**93.59**	4.94	73.53	86.42	83.67	15.24	9.83
3	100	87.65	93.68	93.15	5.11	72.82	86.15	82.97	15.53	10.09
4*	**150**	88.20	94.01	93.56	**4.90**	**74.01**	**86.59**	**83.94**	**14.97**	**9.30**
5	200	87.10	93.29	92.80	5.46	72.44	85.41	82.80	16.01	10.30

3.3.2 Optimizing the penalty parameter w^* in the OS-loss function

We also optimize the penalty parameter w^* in Equation (4) as in Table 3. The shape constraint module is added from the 150-th training epoch here. It can be seen from this table that the best performance of the penalty parameter w^* is achieved at $w^* = 1.0$ and all the metrics are consistently the best one among different paramters.

Table 3: Parameter optimization of the penalty weight w^* for the ResU-net plus OS-loss. The training set is S1-training (525 photographs) and the validation set is W10K-validation.

No.	w^*	Optic Disc				Optic Cup				E_{CDR}
		IoU	F1	Dice	SE	IoU	F1	Dice	SE	
1	0.8	86.35	92.80	92.35	5.65	71.71	84.73	82.43	16.64	10.67
2	0.9	87.18	93.39	92.84	5.23	70.72	83.98	81.47	17.46	11.58
3*	**1.0**	**88.20**	**94.01**	**93.56**	**4.90**	**74.01**	**86.59**	**83.94**	**14.97**	**9.30**
4	1.1	87.96	93.84	93.38	5.15	73.59	86.48	83.67	15.11	9.80
5	1.2	86.91	93.25	92.72	6.06	71.4	84.96	82.03	16.43	11.07

3.3.3 Evaluation of the OS-loss Framework on Two Datasets

First, for the optic disc ROI localization, the accuracy of 100% is achieved in both datasets.

(1) On the W10K dataset: To evaluate our proposed OS-loss network, we train the baseline U-net[14] , ResU-net [16] and CE-net [6] from S1 and S2 individually. The S1 represents the case with relatively limited training data, and the S2 represents the case with

sufficient training data. Since the medical image annotation is extremely expensive, this S2 case is often not available in real-world applications. As can be seen from Table 4 (trained from S1) and Table 5 (trained from S2), our proposed OS-loss function can significantly improve the performance consistently in terms of IoU, F1, Dice, SE, and E_{CDR} in most cases, with OS vs. without OS. The CE-net+OS setting achieves the best performance, compared to U-net [14] and ResU-net [16] with the OS-loss function.

Table 4: Performance(%) comparisons of different methods for segmenting OD and OC. The training set is **S1(525)** and the test set is **W10K-testing (1,482 fundus photographs)**.

Methods	Optic Disc				Optic Cup				E_{CDR}
	IoU	F1	Dice	SE	IoU	F1	Dice	SE	
U-net[14]	86.20	92.71	92.34	6.30	69.46	82.93	80.81	18.7	10.34
U-net+OS	86.61	93.09	92.58	5.93	71.35	85.21	81.92	18.06	10.32
ResU-net[16]	86.80	93.12	92.71	5.91	70.31	83.44	81.52	17.98	10.33
ResU-net+OS	87.37	93.48	93.06	5.66	73.32	86.24	83.54	**15.58**	9.30
CE-net[6]	86.36	93.04	92.37	5.48	73.02	85.96	83.24	16.18	10.99
CE-net+OS	**88.69**	**94.36**	**93.85**	**4.75**	**74.50**	**87.34**	**84.26**	15.79	**8.89**

Table 5: Performance(%) comparisons of different methods for segmenting OD and OC. The training set is **S2(6,391)**, and the test set is **W10K-testing (1,482 fundus photographs)**.

Methods	Optic Disc				Optic Cup				E_{CDR}
	IoU	F1	Dice	SE	IoU	F1	Dice	SE	
U-net[14]	90.43	95.26	94.93	3.41	77.95	88.02	86.94	13.53	6.27
U-net+OS	90.19	95.13	94.79	3.27	80.71	90.43	88.73	11.21	5.88
ResU-net[16]	90.66	95.24	95.04	3.29	78.92	89.20	87.60	12.58	6.17
ResU-net+OS	90.52	95.16	94.97	3.41	82.19	91.37	89.63	10.29	5.55
CE-net[6]	90.35	95.19	94.88	3.04	82.04	91.50	89.62	10.62	5.43
CE-net+OS	**91.04**	**95.55**	**95.27**	**2.83**	**83.34**	**91.96**	**90.45**	**9.74**	**5.32**

Further, we offer several typical examples to prove the effectiveness of our method, as in Figure 6. The CDR of the ground truth ranges from 25.95% to 91.52% and the actual size of the optic disc or cup varies among different examples. As in Figure 6(4)(6)(8), the predicted optic disc contours of the original ResU-net trained on S1 have strong geometric distortions with relatively sharp corners and non-convexity shape. For other examples, these distortions become weaker but there are still the optic cup contours having unacceptable parts. The ResU-net using our OS-loss function trained on S1 can improve these results significantly.

When training based on S2 with relatively sufficient data, the performance of the original ResU-net improves significantly compared to the same model trained on S1. In this case, the ResU-net using the OS-loss function does not have significant improvements as the current baseline has achieved relatively good segmentation results, resulting in a limited improving space. Besides, compared to ResU-net+OS(S1) to ResU-net(S2), our method with limited training data still can get comparative results.

From these results, conclusions can be made that our OS-loss framework can achieve reasonable shape contours, which can avoid the geometric distortions in most cases.

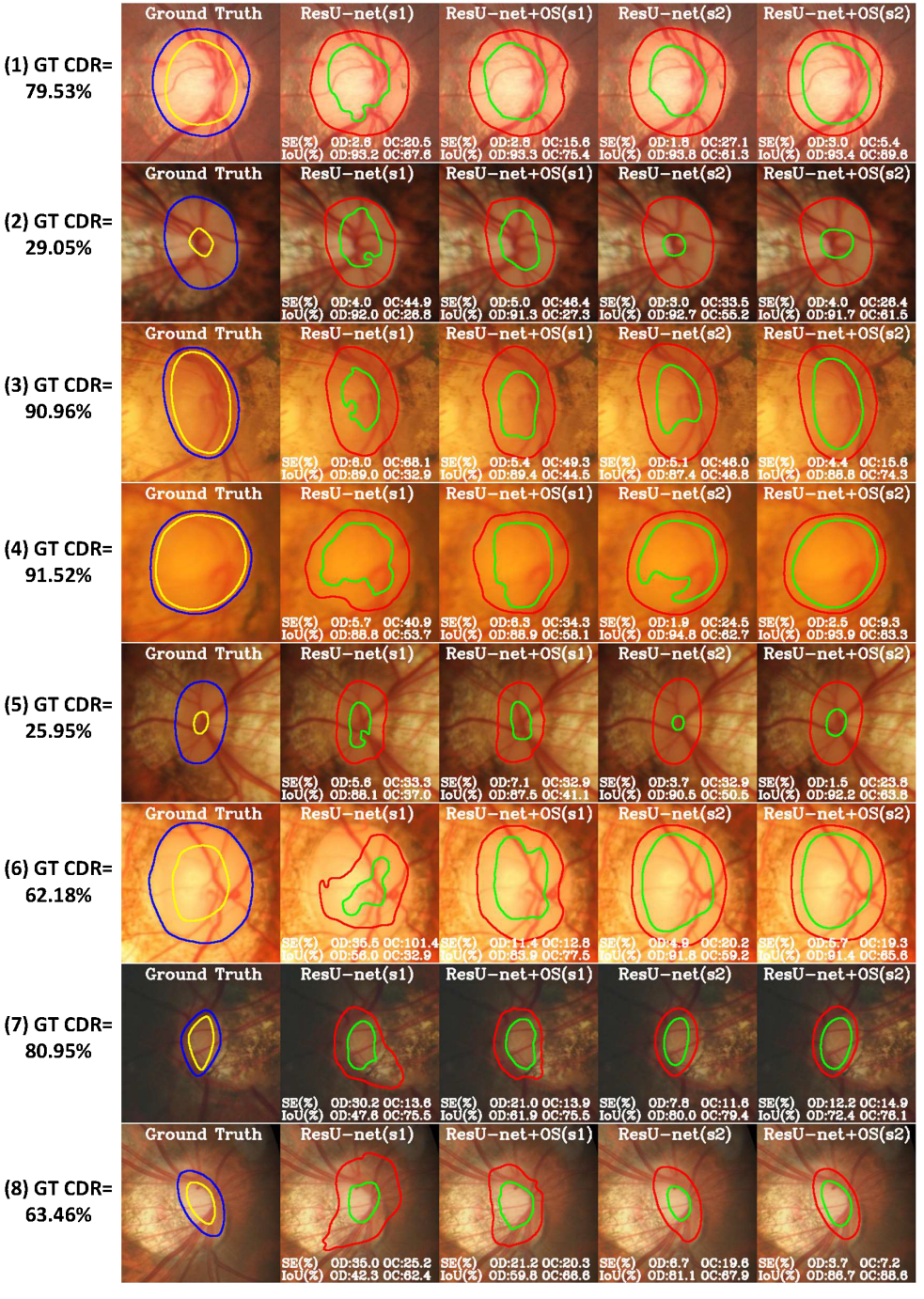

Figure 6: Predicted Results of the optic disc and cup segmentation on the W10K dataset.

(2) On the RIM-ONE-r3 dataset: We also evaluate our method on the public RIM-ONE-r3 dataset. As shown in Table 6, the method using the OS-loss function outperforms the baseline networks (U-net [14] , ResU-net [16] , and CE-net[6]) in most cases, except for the best IoU and Dice for the optic disc using U-net and the best IoU for the optic cup using ResU-net. Due to the small data size and the lacking cases of obvious geometric distortions, three baseline networks already have relatively high performances, and our OS-loss function does not have significant improvements for three comparing groups (using OS-loss or not, *i.e.,* line 2 vs. 1, line 4 vs. 3, line 6 vs. 5). Compared to the original CE-net, the mean CDR error of the CE-net+OS method in the RIM-ONE-r3 dataset decreases around 1.98%, and IoU and Dice for the optic disc and cup improve significantly. There is also a similar case for the ResU-net, except for a slight increase (0.06%) in terms of the mean CDR error.

Table 6: Performance (%) comparisons of different methods on the public RIM-ONE-r3 [4].

No.	Methods	Optic Disc		Optic Cup		E_{CDR}
		IoU	Dice	IoU	Dice	
1	U-net[14]	**94.63**	**97.23**	86.07	91.60	3.30
2	U-net+OS	94.43	97.12	86.43	91.83	3.55
3	ResU-net[16]	94.34	97.08	**88.00**	91.34	**2.68**
4	ResU-net+OS	94.52	97.17	86.38	91.83	2.74
5	CE-net[6]	93.78	96.76	84.48	91.03	4.94
6	CE-net+OS	94.57	97.19	87.18	**92.38**	2.96

As a summary, the experiments on both the public RIM-ONE-r3 and the private W10K datasets prove that the proposed method using the OS-loss function is better than the original U-net [14] , ResU-net [16] , and CE-net [6] , especially for the baseline network that outputs more unreasonable segmentation results with strong geometric distortions.

4 Conclusions

For the pixel-wise segmentation of the optic disc and cup in fundus photographs, we have proposed a novel loss function, called OS-loss, which sufficiently explores the oval shape constraint into a baseline deep learning network. At a certain training epoch, the loss function will be updated to assign bigger penalty weights for the misleading pixels from the verified penalty point set. Also, an oval-friendly metric called shape error (SE) is proposed to better reflect the fitness of two oval contours. Experiments on the public RIM-ONE-r3 dataset with 159 fundus photographs and a private W10K dataset with 9,879 fundus photographs show that our proposed OS-loss function outperforms the state-of-the-art method.

5 Acknowledgement

This work is supported by the CSC State Scholarship Fund (201806295014), NSFC (No. 61672523, No.61771468), CAMS Initiative for Innovative Medicine (2018-I2M-AI-001), Beijing Natural Science Foundation (No.4192029), Beijing Hospitals Authority Youth Programme (QML20170206); The priming scientific research foundation for the junior researcher in Beijing Tongren Hospital, Capital Medical University (2018-YJJ-ZZL-052).

References

[1] Ahmed Almazroa and et al. Optic disc segmentation for glaucoma screening system using fundus images. *Clinical Ophthalmology*, 11:2017–2029, 2017.

[2] Jun Cheng and et al. Quadratic divergence regularized SVM for optic disc segmentation. *Biomedical Optics Express*, 8(5):2687–2696, 2017.

[3] Yinwei Dong and et al. Interactive cup and disc segmentation of fundus images based on c-v model. *Computer application and software*, 29(5):104–108, 2012.

[4] Francisco Fumero and et al. Interactive tool and database for optic disc and cup segmentation of stereo and monocular retinal fundus images. In *WSCG*, 2015.

[5] Lena Gorelick and et al. Convexity shape prior for segmentation. In *ECCV*, 2014.

[6] Zaiwang Gu and et al. Ce-net: Context encoder network for 2d medical image segmentation. *CoRR*, abs/1903.02740, 2019.

[7] Kaiming He and et al. Deep residual learning for image recognition. In *CVPR*, 2016.

[8] Gopal Datt Joshi and et al. Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment. *IEEE T-MI*, 30(6):1192–1205, 2011.

[9] Jonathan Long and et al. Fully convolutional networks for semantic segmentation. In *CVPR*, pages 3431–3440, 2015.

[10] Diego Marcos and et al. Learning deep structured active contours end-to-end. In *CVPR*, pages 8877–8885, 2018.

[11] Pardha Saradhi Mittapalli and et al. Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. *BSPC*, 24:34–46, 2016.

[12] Hariharan Ravishankar and et al. Learning and incorporating shape models for semantic segmentation. In *MICCAI*. Springer, Cham, 2017.

[13] Shaoqing Ren and et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE T-PAMI*, 39(6):1137–1149, 2017.

[14] Olaf Ronneberger and et al. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015.

[15] Loïc Alain Royer and et al. Convexity shape constraints for image segmentation. In *CVPR*, 2015.

[16] Sharath M Shankaranarayana and et al. Joint optic disc and cup segmentation using fully convolutional and adversarial networks. In *FIFI & OMIA*, pages 168–176, 2017.

[17] Song Wang and et al. Global detection of salient convex boundaries. *International Journal of Computer Vision*, 71(3):337–359, 2007.

[18] Shan Zheng and et al. Automatic segmentation of optic disc and cup using multiphase active contour model in fundus images. *Journal of Image & Graphics*, 2014.