

# Optic Cup Segmentation Using Large Pixel Patch Based CNNs

Yundi Guo<sup>1,2</sup>, Beiji Zou<sup>1,3</sup>, Zailiang Chen<sup>1,2</sup>\*, Qi He<sup>1,2</sup>, Qing Liu<sup>1,2</sup>, and Rongchang Zhao<sup>1,2</sup>

<sup>1</sup> School of Information Science and Engineering, Central South University, Changsha, China

<sup>2</sup> Center for Ophthalmic Imaging Research, Central South University, Changsha, China

<sup>3</sup> Center for Information and Automation of China Nonferrous Metals Industry Association, Changsha, China

**Abstract.** Optic cup(OC) segmentation on color fundus image is essential for the calculation of cup-to-disk ratio and fundus morphological analysis, which are very important references in the diagnosis of glaucoma. In this paper we proposed an OC segmentation method using convolutional neural networks(CNNs) to learn from big size patch belong to each pixel. The segmentation result is achieved by classification of each pixel patch and postprocessing. With large pixel patch, the network could learn more global information around each pixel and make a better judgement during classification. We tested this method on public dataset Drishti-GS and achieved average F-Score of 93.73% and average overlapping error of 12.25%, which is better than state-of-the-art algorithms. This method could be used for fundus morphological analysis, and could also be employed to other medical image segmentation works which the boundary of the target area is fuzzy.

**Keywords:** Optic Cup, CNNs, Pixel Patch, Segmentation

## 1 Introduction

Accurate optic cup(OC) segmentation on color fundus image is important for the diagnosis of glaucoma. Glaucoma is the world's first leading cause of irreversible blindness and has grown up to be a global public health problem. The clinical manifestation of glaucoma is mainly optic atrophy. The optic nerve head which usually called the optic disk(OD) passes visual information from the retina to the brain. The OD has a vertical elliptical shape[19] and is divided into two separate zones: the optic cup and the neuroretinal rim. The atrophy changes the shape of the OC region. Therefore, analyzing the shape of the OC and OD quantitatively is crucial.

There are many works for OD segmentation and the precision is very high now[15][1][14]. Unlike the OD, good methods for OC segmentation are still scarce

---

\* Corresponding author

and the existing results are low in precision that cannot support the diagnosis of glaucoma well. The boundary of OC is not as clear as OD, and the vascular cross point is always on it, so that OC segmentation is much harder in both manually and automatically. Some methods tried to segment OC with the detection of blood vessel[20] kinks because the 3D structure of OC makes the vessel bend at the boundary in 2D fundus images[4][3]. The challenge is to find a good initial estimation of the cup boundary and exclude vessel bends from a non-cup boundary. Statistical model-based methods tried to use regional color characteristics to find the best or an adaptive threshold to fit a boundary of the OC[18][6]. But in many fundus images, there is no obvious color difference within the disk to mark the cup boundary. Jun Cheng[2] proposed OD and OC segmentation using superpixel classification. Mittapalli[11] proposed a fuzzy c-means(FCM) based method for cup segmentation. Pixel clustering based methods could make an approximate boundary of the OC, but segmenting the OC accurately is still a challenging problem.

Deep convolutional neural networks (CNNs)[10][5] are a type of multi-layer, fully trainable models that can capture highly nonlinear mappings between inputs and outputs. These models were originally motivated from computer vision problems and thus are intrinsically suitable for image-related applications. Since the early 2000s, CNNs have been applied with great success to the detection, segmentation and recognition of objects and regions in images[9].

In this paper, we proposed an OC segmentation method using CNNs with large pixel patch to train and predict pixels belong to the OC area. Using large pixel patch, more global information around each pixel is learned by the network. With two-dimensional convolution and non-linear layers, the color features, vessel kinks or other unknown characteristics of the OC could be highly abstractly extracted. We tested this method on public dataset Drishti-GS and achieved average F-Score of 93.73% and average overlapping error of 12.25%, which is better than state-of-the-art methods. This method could also be employed to other medical image segmentation works which the boundary of the target area is fuzzy.

## 2 Method

We use CNNs to automatically extract features around each pixel through a large patch. The principle of the convolutional neural networks is like people's eyes and brains to deal with what they see[8]. With large pixel patch, the networks could get as much as the global perception around each pixel.

### 2.1 OD Area Cut and Shading Unification

At the start we cut the square area which the OD relatively centers to, which is the region of interest[21]. This operation allows us to concentrate on OC segmentation without considering much useless background. There are numerous mature research methods for OD detection like the ASM-based method[17],

method based on approximate nearest neighbour field[12] and so on. As they are not the key point of this work, the details are omitted here.

Because fundus images are not generally taken in consistent illumination intensity, the variation of RGB value in different images is unfavorable for the training. So we convert the cut images from RGB to Lab color space and apply an equalization operation according to Formula 1 to all the three channels.

$$P_i = \begin{cases} P_{mean} + (P_i - P_{mean}) / (P_{max} - P_{mean}) & P_i > P_{mean} \\ P_{mean} - (P_{mean} - P_i) / (P_{mean} - P_{min}) & P_i < P_{mean} \\ P_{mean} & P_i = P_{mean} \end{cases} \quad (1)$$

Where  $P_{mean}$ ,  $P_{max}$ ,  $P_{min}$ ,  $P_i$  means the average value, the maximum value, the minimum value of all the pixel values in the channel and the  $i$ th pixel value respectively.  $i$  denotes the pixel index.

A Lab color space is based on nonlinearly compressed with dimension  $L$  for lightness and  $a$ ,  $b$  for the color-opponent dimensions. This operation will reduce color differences from image to image to a certain level.

## 2.2 Data Augmentation and Training Data Extracting

In the field of medical science, the amount of available data with labels is relatively small to other fields such as natural image processing. In this case, data augmentation is needed to create more meaningful data from finite marked images. For fundus images, there are usually images from the left eye and the right eye and the two are usually not identical. We flip the images horizontally as new training data so that the amount of data with labels is doubled.

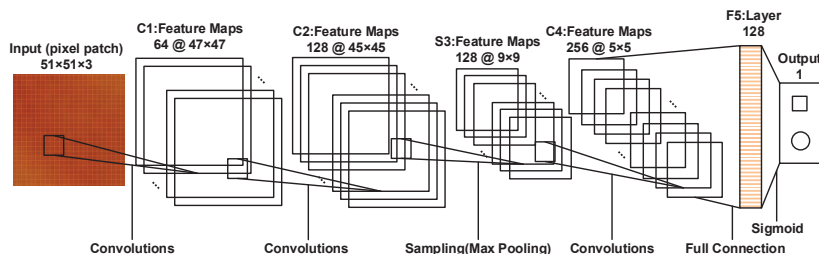
Theoretically, the higher the image resolution is, the more accurate the results can be achieved. But the size of input images is proportional to the amount of computation and memory. To make a balance, each image is scaled down to  $120 \times 120$  to reduce the data scale for our CNN without obvious influence on the results in practice.

Because the boundary of OC is fuzzy, local information like intensity gradient or value is not sufficient to classify a pixel. We extract patches of size  $51 \times 51$  for each pixel as the input data for our CNN. This big size patch covers much area of the whole square so that the network will learn more global information around each pixel. The size we chose is a trade off between sufficient information and computational burden. We also considered input patch sizes of  $41 \times 41$  and  $61 \times 61$  and so on.

## 2.3 Deep CNN Architectures

In this section we describe the architecture of our CNN with input patch size of  $51 \times 51$  to explain the method we use. This CNN architecture contains three input feature maps corresponding to RGB channels. There are three convolutional layers, one max-pooling layer and one full connection layer in the network. Each layer is followed with a nonlinear filtering activation function. 0.25 dropout is

applied to mitigate overfitting. A sigmoid layer is added to the end for binary classification. The structure is shown in Figure 1.



**Fig. 1.** CNN structure for OC pixel training and prediction. The net of CNN contains 5 layers: three convolutional layers, one max-pooling layer and one full connection layer.  $51 \times 51 \times 3$  denotes the dimension of the input image.  $64 @ 47 \times 47$  denotes 64 feature maps of size  $47 \times 47$  in that layer.

## 2.4 Postprocessing

Pixel based prediction has a shortcoming that the predicted OC area is composed of isolated points. Some points may be apart from the unicom region and the shape of the predicted area may be not regular. Most of the isolated points are caused by prediction error. We use erosion and dilation to eliminate these isolated points. After that, as the OC area is usually like an ellipse, ellipse fitting operation is utilized to the result pixels.

## 3 Experiments

We use the public data set Drishti-GS[13] to evaluate our method. Though there are many fundus image datasets, most of them are not available online or publicly. Other available datasets do not incorporate challenges of OD and OC segmentation. Drishti-GS is a dataset of retinal images for OD and OC segmentation along with ground truth. There are 50 images can be directly downloaded. Open source library Keras is used for the construction of the networks. GTX TitanX $\times 4$  GPU speed-up workstation is deployed to perform the experiments. We will firstly test the performance of different patch sizes as input, and then compare our method with several other methods.

### 3.1 Performance on Different Patch Sizes

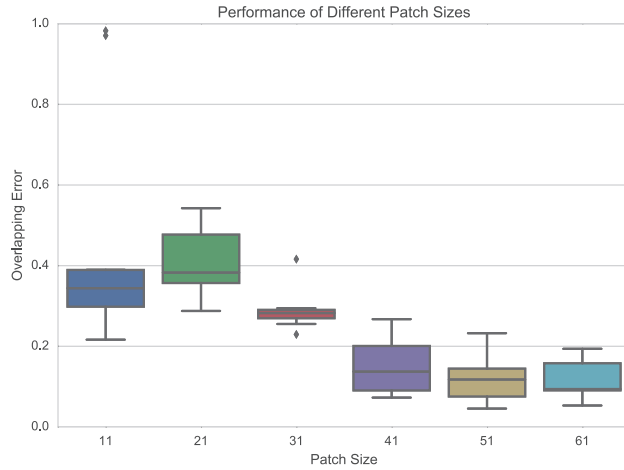
Different patch sizes as input are tested by CNNs. We use 40 images as training data and 10 as testing data. The overlapping error  $OE$  is used as the evaluation criterion. It is calculated by Formula 2.

$$OE = 1 - \frac{||S \cap G||}{||S \cup G||} \quad (2)$$

Where  $S$  is the segmentation result,  $G$  is the ground truth, and  $||S \cap G||$  is the number of shared positive pixels by  $S$  and  $G$ ,  $||S \cup G||$  is the number of positive pixels in  $S$  or  $G$ . The overlapping error lies in  $[0, 1]$  and smaller value indicates a higher segmentation accuracy. The average training time of each patch size is shown in Table 1. The overall overlapping error of each patch size is shown by boxplot in Figure 2.

**Table 1.** Average training time(ATE) of different patch sizes for OC segmentation

Patch Size	11×11	21×21	31×31	41×41	51×51	61×61
ATE(s/image)	5	11	32	73	95	113



**Fig. 2.** Box plot of the segmentation performance achieved by CNNs over 6 different patch sizes. The plot uses overlapping error to measure the performance for different patch sizes. The patch sizes are  $11 \times 11$ ,  $21 \times 21$ ,  $31 \times 31$ ,  $41 \times 41$ ,  $51 \times 51$  and  $61 \times 61$  respectively. The central mark represents the median, the edges of the box denote the 25% and 75% in result order.

With the increasing of patch size, the results become better and stable. With small patch, the network can only see a tiny scope around each pixel. The color does not change too much in the limited area and other pixel patches not belong to OC may be quite similar to the ones in OC area, so the network will be easily

overfitted during training. When the patch size becomes larger, more global information is contained. However, with bigger patch size, more computer memory is needed as well. In our experiments, to achieve better performance, patch size of  $61 \times 61$  is acceptable. The prediction phase is fast and the computing time can be ignored.

### 3.2 Comparison With Other Methods

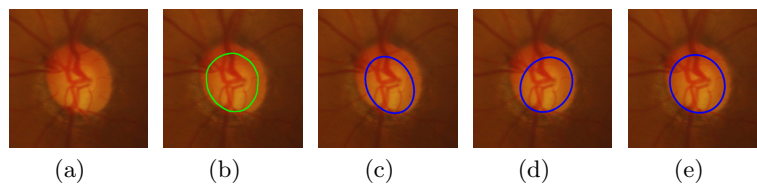
There are several previous methods have been proposed these years. Methods like threshold[7], r-bend[7], ASM[18], regression[16] are well compared in the latest papers like [2], [11] and so on. So we conducted experiments to compare with the superpixel based method[2] and the spatially weighted fuzzy c-means(SWFCM) based method[11]. We use the results of patch size  $61 \times 61$  for the comparison, of which the test score is 0.35 and the test accuracy is 0.92. For better quantitative evaluation, we add *F-Score* to measure the segmentation results according to ground truth. They are defined as Formula 3.

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}, F\text{-Score} = \frac{2 \times P \times R}{P + R} \quad (3)$$

Where *TP*, *FP*, *FN* means true positives, false positives and false negatives of the classified pixels respectively. *P* is precision and *R* is recall. The *P* and *R* are calculated based on the overlapped area between the detected contour and the groundtruth.

**Table 2.** Comparison of different methods for OC segmentation

Method	<i>OE</i>	<i>F-Score</i>
Superpixel[2]	24.31%	85.98%
SWFCM[11]	21.21%	87.91%
Proposed	12.25%	93.73%



**Fig. 3.** Sample results. From left to right columns: (a) the original images, (b) the ground truth, (c) is by superpixel[2], (d) is by SWFCM[11], and (e) is the proposed method.

Table 2 and Figure 3 shows the results according to ground truth. The proposed method achieves an overlapping-error of 12.25% and *F-Score* of 93.73,

which means that the segmented cups are mostly overlapped with the ground truth. This enhanced performance is mainly because it combined the around area together of each pixel to automatically learn the image features. Each pixel whether belong to OC is deeply analysed like how people's eye and brain work. Methods only considering local color or manually setting related features are weak to tackle with complex OC area. When the dataset is small, the generalization ability of the proposed CNN based method is better than other machine learning based methods. Because of the big size patch and CNN, which can peer the global features automatically, the grey level and the blood vessel shape are effectively combined together automatically without special design.

## 4 Conclusion

In this paper, we proposed an OC segmentation method using CNNs with large pixel patch to train and predict pixels belong to the OC area. With the help of GTX TitanX×4 GPU speed-up workstation, big size patch is allowed to train and the speed becomes acceptable. The CNNs make full use of the global information around each pixel and output the prediction of pixel labels. We tested this method on different pixel patch size and confirmed that bigger pixel patch makes better segmentation. Results on the Drishti-GS dataset showed that our proposed model significantly outperformed prior methods on OC segmentation. In this work, the OC segmentation problem is formulated as patch classification task that much duplicate information is reprocessed. In the future work we will improve the model to combine the whole image together for pixel prediction.

## References

1. Aquino, A., Gegúndez-Arias, M.E., Marín, D.: Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *Medical Imaging, IEEE Transactions on* 29(11), 1860–1869 (2010)
2. Cheng, J., Liu, J., Xu, Y., Yin, F., Wong, D.W.K., Tan, N.M., Tao, D., Cheng, C.Y., Aung, T., Wong, T.Y.: Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *Medical Imaging, IEEE Transactions on* 32(6), 1019–1032 (2013)
3. Fondon, I., Valverde, J.F., Sarmiento, A., Abbas, Q., Jimenez, S., Alemany, P.: Automatic optic cup segmentation algorithm for retinal fundus images based on random forest classifier. In: *EUROCON 2015-International Conference on Computer as a Tool (EUROCON)*, IEEE. pp. 1–6. IEEE (2015)
4. Hatanaka, Y., Nagahata, Y., Muramatsu, C., Okumura, S., Ogohara, K., Sawada, A., Ishida, K., Yamamoto, T., Fujita, H.: Improved automated optic cup segmentation based on detection of blood vessel bends in retinal fundus images. In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. pp. 126–129. IEEE (2014)
5. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580* (2012)

6. Issac, A., Parthasarathi, M., Dutta, M.K.: An adaptive threshold based algorithm for optic disc and cup segmentation in fundus images. In: *Signal Processing and Integrated Networks (SPIN)*, 2015 2nd International Conference on. pp. 143–147. IEEE (2015)
7. Joshi, G.D., Sivaswamy, J., Karan, K., Krishnadas, S.: Optic disk and cup boundary detection using regional information. In: *Biomedical Imaging: From Nano to Macro*, 2010 IEEE International Symposium on. pp. 948–951. IEEE (2010)
8. LeCun, Y., Bengio, Y.: Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks* 3361(10), 1995 (1995)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* 521(7553), 436–444 (2015)
10. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11), 2278–2324 (1998)
11. Mittapalli, P.S., Kande, G.B.: Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. *Biomedical Signal Processing and Control* 24, 34–46 (2016)
12. Ramakanth, S.A., Babu, R.V.: Approximate nearest neighbour field based optic disk detection. *Computerized medical imaging and graphics* 38(1), 49–56 (2014)
13. Sivaswamy, J., Krishnadas, S., Datt Joshi, G., Jain, M., Ujjwaft Syed Tabish, A.: Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In: *Biomedical Imaging (ISBI)*, 2014 IEEE 11th International Symposium on. pp. 53–56. IEEE (2014)
14. Tjandrasa, H., Wijayanti, A., Suciati, N.: Segmentation of the retinal optic nerve head using hough transform and active contour models. *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 10(3) (2012)
15. Welfer, D., Scharcanski, J., Kitamura, C.M., Dal Pizzol, M.M., Ludwig, L.W., Marinho, D.R.: Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach. *Computers in Biology and Medicine* 40(2), 124–137 (2010)
16. Xu, Y., Xu, D., Lin, S., Liu, J., Cheng, J., Cheung, C.Y., Aung, T., Wong, T.Y.: Sliding window and regression based cup detection in digital fundus images for glaucoma diagnosis. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*, pp. 1–8. Springer (2011)
17. Yin, F., Liu, J., Ong, S.H., Sun, Y., Wong, D.W., Tan, N.M., Cheung, C., Baskaran, M., Aung, T., Wong, T.Y.: Model-based optic nerve head segmentation on retinal fundus images. In: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. pp. 2626–2629. IEEE (2011)
18. Yin, F., Liu, J., Wong, D.W.K., Tan, N.M., Cheung, C., Baskaran, M., Aung, T., Wong, T.Y.: Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis. In: *Computer-based medical systems (CBMS)*, 2012 25th international symposium on. pp. 1–6. IEEE (2012)
19. Youssif, A.A.H.A.R., Ghalwash, A.Z., Ghoneim, A.A.S.A.R.: Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter. *Medical Imaging, IEEE Transactions on* 27(1), 11–18 (2008)
20. Zhu, C., Zou, B., Zhao, R., Cui, J., Duan, X., Chen, Z., Liang, Y.: Retinal vessel segmentation in colour fundus images using extreme learning machine. *Computerized Medical Imaging and Graphics* (2016)
21. Zou, B., Liu, Q., Chen, Z., Fu, H., Zhu, C.: Surroundedness based multiscale saliency detection. *Journal of Visual Communication and Image Representation* 33, 378–388 (2015)