



Large-Scale Left and Right Eye Classification in Retinal Images

Peng Liu^{1,2}(✉), Zaiwang Gu^{2,3}, Fan Liu¹, Yuming Jiang¹, Shanshan Jiang²,
Haoyu Mao², Jun Cheng², Lixin Duan¹, and Jiang Liu²

¹ Big Data Research Center,
University of Electronic Science and Technology of China, Chengdu, China
liupengimed@nimte.ac.cn

² Cixi Institute of Biomedical Engineering,
Ningbo Institute of Materials Technology and Engineering, Ningbo, China

³ School of Mechatronic Engineering and Automation, Shanghai University,
Shanghai, China

Abstract. Left and right eye information is an important priori for automatic retinal fundus image analysis. However, such information is often not available or even wrongly provided in many datasets. In this work, we spend a considerable amount of efforts in manually annotating the left and right eyes from the large-scale Kaggle Diabetic Retinopathy dataset consisting of 88,702 fundus images, based on our developed online labeling system. With the newly annotated large-scale dataset, we also train classification models based on convolutional neural networks to discriminate left and right eyes in fundus images. As experimentally evaluated on the Kaggle and Origa dataset, our trained deep learning models achieve 99.90% and 99.23% in term of classification accuracy, respectively, which can be considered for practical use.

Keywords: Left and right eye classification
Convolutional neural networks

1 Introduction

Retinal fundus images contain rich information for ophthalmic disease diagnosis. Ophthalmologists often determine the health conditions of eyes by examining blood vessels, optic nerve head, vitreous and macula on the corresponding retinal fundus images. Among those, information on left and right eyes is often used in ophthalmic disease diagnosis. For instance, such information is used to determine the nasal and temporal side of an eye. Also, left and right eye information is also considered in the glaucoma diagnosis, when comparing asymmetric cup to disc ratios of different eyes. Moreover, for the diagnosis of age-related macula degeneration, eye side information is also used to determine the location of the macula.

Although some fundus cameras automatically record left and right eye information when taking the retinal images, many others still do not record such information. In our practical study, we often find that clinicians put all the retinal fundus images from one subject into one folder without having the labeling information of either left or right eyes. Moreover, we have found that some images are inverted (rotated for 180°), which can affect the visual appearance of left vs. right. According to the description of Kaggle Diabetic Retinopathy dataset (Kaggle DR dataset) [4], the retinal images provided are a mix of images shown as standard retina anatomy and images taken through a microscope condensing lens (which are inverted). Figure 1 shows two pairs of retinal images from the dataset. The above cases can actually pose big problems when later eye side information gets necessary needs, which requires immediate actions.

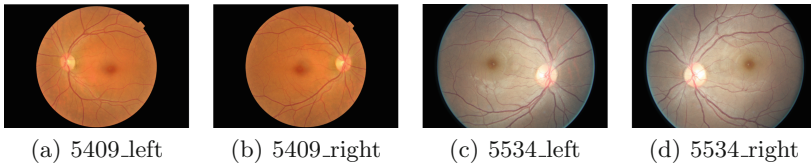


Fig. 1. Examples of retinal images of a pair of eyes not inverted and an pair of eyes inverted in Kaggle DR dataset. (a) and (b) A pair of retinal fundus images from a patient not inverted; (c) and (d) A pair of inverted retinal fundus images from a patient.

In recent years, methods have been proposed to classify left and right eyes. In [9], Tan et al. proposed a classification method by examining the intensity changes across the optic disc. In [8], support vector machine was used to train a more robust classification model. In [7], the vessel distribution within the optic disc was further used to distinguish the left and right eyes in retinal fundus images and tested on the Origa dataset [10]. Those methods have two major limitations. On one hand, they are all built on holistic features (e.g., intensity changes or vessel changes within the disc) and shallow models (e.g., SVM), and are thus sensitive to images of different quality (e.g., taken from different machines) as well as not fully capable of capturing the high semantic meaning based on the image content. On the other hand, the datasets used in their experiments are of small scales. To solve their limitations, in this work we propose to employ deep learning based methods to train better classification models for left and right eye classification. And also, we relabel the large-scale Kaggle DR dataset, which consists of 88,702 fundus images, by providing the information of left and right eyes. To the best of our knowledge, this dataset is so far the largest one for left and right eye classification.

The rest of the paper is organized as follows. Section 2 introduces our online system for left and right eye annotation. Section 3 details the procedures to train deep learning models for left and right eye classification. Experimental results are provided in Sect. 4. Finally, conclusive remarks are presented in Sect. 5.

2 Labeling Protocol

In this paper, we spend a considerable amount of efforts in providing the left and right eye information on the Kaggle DR dataset which is public available¹. This large-scale dataset contains 88,702 fundus images in total: 35,126 training images and 53,576 test images which are already split and provided. Base on the filename, each image is named by a patient id followed by left or right eye information, for example, “5409_left” denotes left eye of patient 5409. However, we found that many images were inverted. According to our statistics, more than 36% images are inverted or contain wrong labels. This surprising number actually motivates us to develop an online system for left and right eye labeling.

2.1 Manual Labeling of Left and Right Eyes and Inverted Images

The left and right eye information can be determined by comparing the location of optic disc with that of macula. If the optic disc on the left side of the macula, the retinal image is from a left eye, otherwise, right eye. However, in some images, the macula is not captured or its location cannot be easily identified due to pathological changes. Therefore, the above method cannot be used. A second method is to examine the intensity changes within the optic disc. Typically, the temporal side of the optic disc is brighter than the nasal side. A third method is to use the blood vessels. Very often, the main vessels bend toward the macula. Simultaneously, we label whether the retinal image is inverted via examining if there is a notch on the side of the image.

In our manual labeling process, we combine the above rules to determine the left and right eye information.

2.2 Online Labeling System

For the ease of manual labeling, we have developed a labeling system to relabel all the 88,702 fundus images from the Kaggle DR dataset. A group of six researchers has been trained to identify left and right eyes and whether the retinal image is inverted. In our relabeling, we label each image as left, right or unable to tell. We label each image as inverted or not inverted as while. In the relabeling, each image will be labeled by two researchers independently. When the label result by the two researchers are different, the image will be examined and discussed by a group of at least three researchers to reach a consensus. For images where nobody can tell the left and right eye information, we retain the original label.

3 Left and Right Eye Classification Based on Deep Learning

Convolutional neural networks (CNNs) [3, 5] have achieved superior performance in object classification and detection. We use three typical CNN architectures,

¹ <https://www.kaggle.com/c/diabetic-retinopathy-detection>.

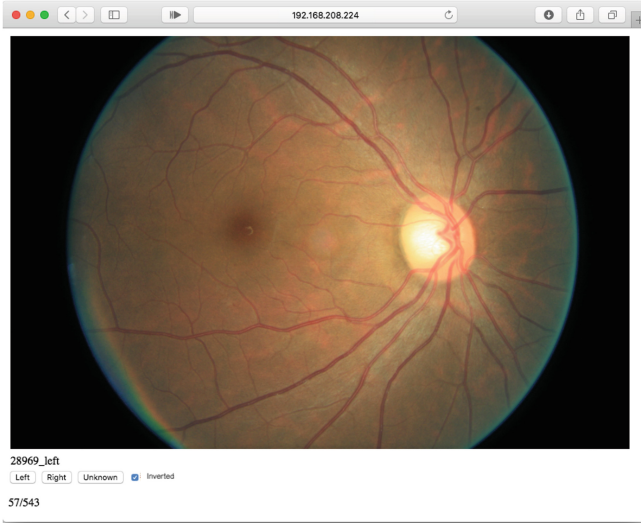


Fig. 2. Examples of the collaboration labeling system. Researchers need to determine left and right eye and whether the image is inverted, then choose “Right” or “Left” to label the retinal fundus image or “Unknown” to indicate a pair of poor quality fundus images.

including the VGG-16 [6], the 50-layer and 101-layer ResNet [3], to automatically determine the left and right eye information (Fig. 2).

3.1 Image Normalization

Since the images in the Kaggle dataset are obtained under different conditions including the age of subjects, the models of the fundus camera and the settings, dilated or un-dilated eyes, illumination changes, these images show a variety change of colors. Very often, image normalization to reduce the changes from one to another image is beneficial for subsequent retinal image analysis.

In this paper, we first preprocessing the images from the Kaggle and Origa datasets. In our processing, we first extract the effective image region by applying a thresholding to the full image to remove the unnecessary black edge, where the threshold is set as 20 of gray value in our implementation.

Next, we normalize the image following the min-pooling’s solution [2]. Mathematically, each image is computed as

$$I_c = \alpha \cdot I + \beta \cdot G(\rho) * I + \gamma, \quad (1)$$

where I represents the input image, $G(\rho)$ denotes the Gaussian filter with a standard deviation of ρ , $*$ means the convolution operation, and α , β , γ are pre-defined parameters (we use $\alpha = 4$, $\beta = -4$, $\rho = 10$, $\gamma = 128$ in the experiments).

In the last step, we resize all the images to 224×224 . For inverted images according to our relabeling, we inverted back the images as retina anatomically (macula on the left, optic nerve on the right for the right eye). Figure 3 shows two sample images before and after our normalization. After normalization, the color difference between the two images is reduced.

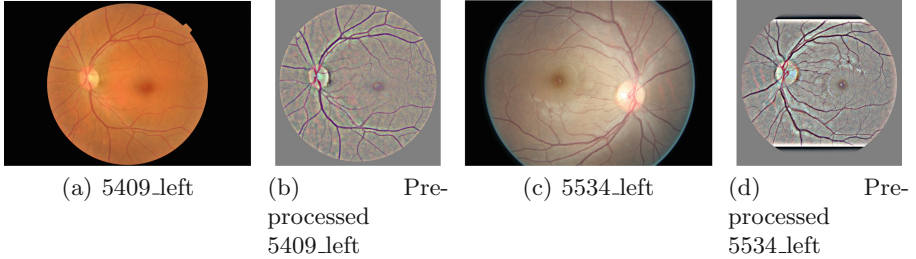


Fig. 3. Two sample images before and after our pre-processing.

3.2 Deep Learning Models

We train the deep learning models from pre-trained networks. Three different CNNs architectures, including VGG-16, 50-layer and 101-layer ResNet, are used in this paper. Since we need to classify left and right eye instead of 1000 classes of objects, we add a 2-d fully connected layer before the softmax layer in the three original architectures.

In the training, we use 35,126 images in the training set of Kaggle DR dataset after our pre-processing with relabeled left and right eye information. We fine-tune the network from the pre-trained using the ImageNet dataset [1]. All the parameters were involved into the finetune. We use Adam optimizer with a learning rate of 0.0001. The models are optimized for 40 epochs and the mini-batch size is 128.

4 Experiments

4.1 Datasets

Kaggle Diabetic Retinopathy Dataset: we utilize the relabeled Kaggle DR dataset which has in total 88,702 retinal fundus images. It was divided into a training set with 35,126 images and a test set with 53,576 images, which is exactly the same as the original partition. According to the our relabeled left and right eye, there are 17,559 left eyes and 17,567 right eyes in the training set, 26,742 left eyes and 26,834 right eyes in the test set.

Origa Dataset: Origa dataset, consists 336 left eyes and 314 right eyes retinal images. In this paper, we applied the finetuned models to classify left and right eyes on all the 650 retinal fundus images.

4.2 Results

Labeling Results. We have labeled left and right eye and inverted information for all the 88,702 fundus images from the Kaggle dataset. Based on our statistics, a total number of 32,199 images have been inverted or contain wrong labels in the original dataset. The detailed statistics are provided in Table 1. It is quite clear that more than 36% images are inverted or provided wrong left and right eye information.

Table 1. Statistics of the left and right eye information on the original Kaggle DR dataset.

Subset	Left/right eye	# Modified labels	# total Images	Modified rate
Training	Left	6,397	17,559	36.43%
Training	Right	6,393	17,567	36.39%
Test	Left	9,754	26,742	36.47%
Test	Right	9,708	26,834	36.18%

Left and Right Eye Classification. We evaluate the finetuned VGG-16, 50-layer and 101-layer ResNet on our newly labeled Kaggle DR dataset and the Origa datasets. We summarize the classification accuracies on the two datasets in Tables 2 and 3, respectively. We can see from the results that image normalization using the Gaussian filter improve the classification performance in all the settings. The best result obtained by ResNet with image normalization indicates its effectiveness and potential in practical use.

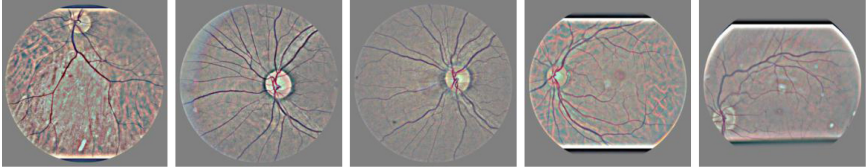
Table 2. Classification accuracies of different models on the Kaggle DR dataset.

Model	Accuracy without Gaussian	Accuracy with Gaussian
ResNet-50	99.53%	99.90%
ResNet-101	99.37%	99.69%
VGG-16	99.20%	99.49%

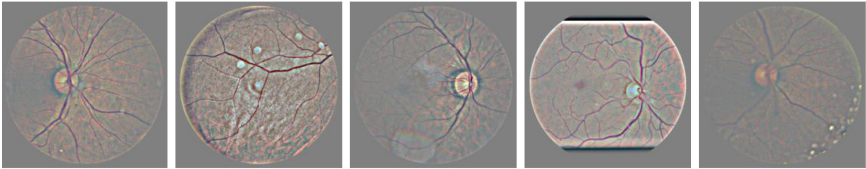
We also present sample images that are both correctly and incorrectly predicted by ReNet-50 in Fig. 4. It is worth noting that the incorrectly predicted images are quite hard to classify even for human experts.

Table 3. Classification accuracies of different models on the Origa dataset.

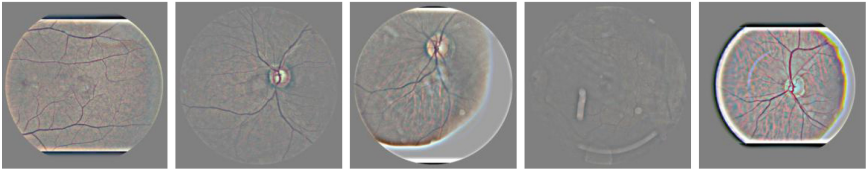
Model	Accuracy without Gaussian	Accuracy with Gaussian
Tan et al. [7]	98.00%	–
ResNet-50	98.62%	99.23%
ResNet-101	98.31%	99.23%
VGG-16	98.15%	98.92%



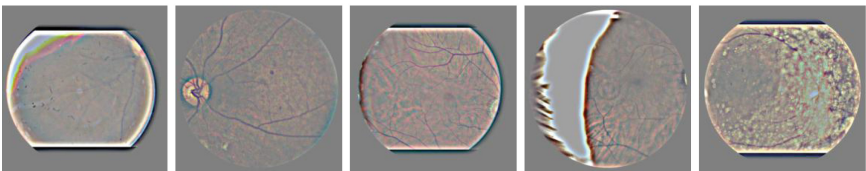
(a) Normalized fundus images containing left eyes, correctly classified.



(b) Normalized fundus images containing right eyes, correctly classified.



(c) Normalized fundus images containing left eyes, incorrectly classified.



(d) Normalized fundus images containing right eyes, incorrectly classified.

Fig. 4. Sample images categorized by the classification results of ResNet-50.

5 Conclusion

In this work, we newly annotate the left and right eye and inverted information for all the 88,702 fundus images from the Kaggle Diabetic Retinopathy dataset. Based on such newly annotated large-scale dataset, we train three CNN models for left and right eye classification from the Kaggle dataset, and evaluate them

on the additional Origa dataset. Extensive experiments clearly show the good generalization ability of the deep learning models as well as their great potential in applying those models for practical use.

References

1. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 248–255. IEEE (2009)
2. Graham, B.: Kaggle diabetic retinopathy detection competition report. University of Warwick (2015)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
4. Kaggle: Diabetic retinopathy detection (2015). <https://www.kaggle.com/c/diabetic-retinopathy-detection>. Accessed 4 Apr 2016
5. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
6. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)
7. Tan, N.M., et al.: An evaluation for an automated left and right eye identification system for digital fundus images for glaucoma diagnosis. *Invest. Ophthalmol. Vis. Sci.* **53**(14), 649–649 (2012)
8. Tan, N.M., et al.: Classification of left and right eye retinal images. *Proc. SPIE* **7624**, 762438 (2010)
9. Tan, N.M., et al.: Automatic detection of left and right eye in retinal fundus images. In: Lim, C.T., Goh, J.C.H. (eds.) 13th International Conference on Biomedical Engineering, pp. 610–614. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-92841-6_150
10. Zhang, Z., et al.: Origa-light: an online retinal fundus image database for glaucoma analysis and research. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3065–3068. IEEE (2010)