Automatic Classification of Cataract based on Deep Learnig

Jingchao Sun School of Software Engineering Beijing University of Technology, Beijing, China Lu Liu School of Software Engineering Beijing University of Technology, Beijing, China

ABSTRACT

Cataract is a serious eye disease which may cause blindness. Early detection is of high significance to the treatment of cataract, which reduces the risk of patients to turning into blindness. Some studies were conducted for fundus image classification based on traditional machine learning methods. However, their performance still can be improved. Therefore, a novel deep learning based method is proposed to classify cataract using fundus images. To avoid relying on mass labelled data, the proposed method resorts to deep transfer learning to reduce the number of parameters that need to be trained. Consequently, in the proposed method, the first five convolutional layers of AlexNet are utilized for general feature learning; then three subsequent layers are designed and fine-tuned for the classification of fundus images. The best performance for cataract detection is 95.37% in terms of classification accuracy.

General Terms

Data Mining, Medical Image Classification, Deep Learning

Keywords

Cataract, Deep transfer learning, Fine tuning, Deep Learning

1. INTRODUCTION

Cataract is a serious eye disease in the world. Cataract is often caused by the buildup of protein in the lens, which may lead to blindness. According to recent studies [9, 4], the number of patients with cataract will reach 40 million in 2025. Therefore, researchers have paid more attention to the detection of cataract.

Accordingly, early detection is very important for patients with cataract, which can make a faster recovery. Unfortunately, skilled ophthalmologists are very scarce in developing areas [7]. The patients with cataract in these areas cannot be diagnosed in time. Currently, medical images are the main tools for the detection of cataract, which includes slit-lamp photography, retro-illumination imaging devices and fundus images. However, the first two diagnosis methods are expensive for patients in developing countries. In addition, manual diagnosis of cataract is a complex and boring work for ophthalmologists. Therefore, it is valuable to provide automatic diagnosis service of cataract for patients in these areas.

Recently, fundus images are demonstrated to be effective for the diagnosis of cataract [7, 18, 19].Comparing with the above two methods, the detection of cataract is simpler. The reason is that diagnosis conclusions can be made according to only one image. In addition, ophthalmologists could examine other eye diseases from fundus images. Therefore, fundus images are more suitable for cataract diagnosis in developing areas. Some studies concerning automatic classification of cataract using fundus images have been conducted [7, 19, 20, 22]. However, they were based on traditional machine learning methods to classify fundus images, the performance of which is needed to be improved.

Therefore, a novel deep learning based method is proposed to classify cataract using fundus images. It is well known that deep learning needs a large amount of labelled data to train parameters in networks. However, there is not often enough data for model training in medical domain. Consequently, deep transfer learning is utilized in the proposed method to reduce the number of parameters that need to be trained. In convolutional neural networks (CNN), shallower layers extract general features and deeper layers learn task-dependent features. Therefore, partial networks trained from the source domain are reused for general feature learning. Specifically, the first five convolutional layers of AlexNet are used for general feature learning; then three subsequent layers are designed and fine-tuned for the classification of cataract. By the combination of deep transfer leaning and fine-tuning, state-of-the-art deep learning technology is applied to achieve automatic cataract diagnosis. The rest of the paper is organized as follows. Section 2 presents the related work. Section 3 describes the methodology. Experimental results are disscussed in section 4. Finally, The conclusion of the paper is given in section 5.

2. RELATED WORK

Nowadays, some researchers have proposed some methods for the classification of cataract. Guo et al. classified fundus images based on a discriminant analysis algorithm and traditional image features [7]. Yang et al. used an ensemble learning based approach for the automatic detection of cataract. They utilized three handcrafted features sets to build SVM models and Back Propagation Neural Network models; then, majority voting and stacking were studied for the combination of base classifiers [19]. Yang et al. exploited a traditional method to detect cataract, in which a neural network was trained based on luminance and texture features [20]. The work in [22] detected cataract using linear discriminant analysis and AdaBoost.

Some other methods were also presented to classify cataract using slit-lamps or retro-illumination images. For examples, Li et al. extracted some handcrafted features to build an SVM model for the



Fig. 1: An overview of the proposed method for the classification of cataract.

grading of cataract [13, 12]. Gray level image gradient-based features were employed for automatic grading of nuclear cataract [17]. The work in [6] used an deep feature learning approach for grading nuclear cataract. By retro-illumination images, the spoke-like features of cortical opacity were utilized to detect cortical opacities and grade the severity of cortical cataract in [11]. A linear discriminant analysis model was trained on texture features for cataract diagnosis in [5]. In addition, some models were constructed based on diagnostic data [1, 15].

However, most of these studies used traditional machine learning methods to build the cataract diagnosis models, in which feature extraction was complex. Therefore, we propose a novel deep learning based method to classify cataract using fundus images, which can obtain better prediction performance.

3. METHODOLOGY

3.1 ALexNet

CNN often consist of multiple convolution and pooling layers with two or three fully connected layers.[10, 16]. To handle multi-class classification tasks, the softmax function is used for the final predictions. Its definition is as follows:

$$S_i = \frac{e^i}{\sum_j e^j} \tag{1}$$

where *i* is the index of i-th class for a sample and S_i is the probability of the sample belonging to i-th class; *j* denotes the indexes of classes. CNN is a feature learning based method, which could extract representative features from images. Comparing with hand-crafted features, deep learning extract features automatically, which are better for training effective models. In addition, shallower layers of CNN can learn general characteristics of images. Therefore, we use the first five layers of AlexNet to extract general features from fundus images in the experiment, the process of which is presented in Fig. 1.

AlexNet is a popular deep learning architecture. It has eight layers, which includes five convolutional layers, some of which are followed by max-pooling layers, and three are fully-connected (FC) layers [10]. The output of the last FC layer is utilized to generate the final predictions with the softmax function. The predictions have the same distribution of the classes in training data. In AlexNet, we can find that the last pooling layer outputs exist 256 feature maps with a 6×6 resolution. Then, a 4096-dimensional feature vector is generated through three FC layers, which represents an image for the final classification. Since the FC layers learn task-dependent features, we only reuse the first five layers instead of the whole

architecture for the general feature extraction from fundus images [21].

3.2 Deep Transfer Learning

Transfer learning is a technology that transfers learnt information from the source to the target domain. It is achieved by reusing existing models. Therefore, it provides researchers with a way to build models with less labelled training data. The most advantage of transfer learning is that the it does not require training data and testing to be Independently and Identically Distributed.

It is different to construct an effective deep learning model for cataract classification with less annotated samples. Consequently, deep transfer learning is investigated for fundus image classification. Currently, many classical deep learning models were proposed for the ImageNet dataset [2], which are eligible for deep transfer learning.

In this paper, the objective of deep transfer learning is also that apply knowledge learnt by deep models to the fundus image classification task. We reuse partial network of AlexNet and the corresponding parameters in the proposed method for the general feature learning. Then, three subsequent layers are fine-tuned based the extracted general feature maps to obtain a good classification performance

3.3 The Porposed Method

An overview of the proposed method is shown in figure 1. Accordingly, AlexNet was trained on the ImageNet dataset. We transfer the first five layers of AlexNet for the general feature learning. The parameters in the five layers are frozen. To obtain the model for the classification of cataract, we design three subsequent layers. A skip connection is used to achieve a better generalization [8], which also makes the training faster. Then, to reduce the redundancy of extracted feature maps, global average pooling is used for feature reduction, which results in a 1024-dimensional feature vector for each fundus image. The last three layers are fine-tuned using fundus images. Finally, the softmax function is utilized to make the final predictions.

The training of the proposed methods is as follows. Let f_p represent the mapping function, where p is the set of parameters. Given a dataset D, let $(x_i, y_i) \in D$, i = 1, ..., N denote the training data; Each x_i is a fundus image; y_i is the label of x_i . p can be trained by solving the following problem:

$$\min_{p} \frac{1}{N} \sum_{n=1}^{N} loss(softmax(f_{p}(x_{n})), y_{n})$$
(2)



Fig. 2: An example of a non-cataractous fundus image and its corresponding R/G/B component images.

Table 1. : THE EXPLANATION OF TP, FN, FP AND TN

Confusion Matrix		Predicted	
		Cataract	Non-cataractous
Actual	Cataract	TP	FN
	Non-cataractous	FP	TN

Where *loss* is the cross-entropy loss function. The extreme value of the objective function is obtained by mini-batch stochastic gradient descent and back propagation.

4. EXPERIMENTS

4.1 Setup of Experiments

4.1.1 Dataset. The fundus images in this experiment are obtained from Beijing Tongren Eye Center of Beijing Tongren Hospital, which is a subset from the database of a Picture Archiving and Communication System. It includes 8030 fundus images(4671 non-cataractous images and 3359 cataract images). The age range of patients is from 10 to 90.

In this work, we focus on the classification of cataract, i.e., building a deep model to decides whether a fundus image is a cataract or not. It is significant for large scale cataract screening.

4.1.2 *Image Preprocessing.* Since the sizes of fundus images are not same, image preprocessing is necessary before model construction. All fundus images are resized to 256×256 pixels. In addition, patients' personal medical information is eliminated from fundus images.

Considering the fact that the quality of fundus images is easily impacted by light, the RGB fundus images are converted to greenchannel (G-channel) fundus images, which are demonstrated to holds more valuable information than original RGB images [14, 3]. A fundus image and its R/G/B-channel images are presented in Fig. 2. Compared with the other three images, the G-channel image is clearer.

4.1.3 Evaluation Criteria. In this paper, the proposed method is tested based on evaluation criteria i.e., accuracy (acc), sensitivity (se) and specificity (sp). Their statistical definitions are listed as

follows:

Accuracy =
$$\frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity = $\frac{TP}{TP + FN}$
Specificity = $\frac{TN}{TN + FP}$

Table 1 lists the meaning of TP, FN, FP and TN. TP (TN) means the number of cataract (non-cataractous) which are correctly predicted. FP (FN) represents the number of non-cataractous (cataract) which are classified as cataract (non-cataractous).

4.2 Experimental Results

9

ç

In this study, three experiments are conducted. The first experiment investigates the performance of the proposed method with different number of fine-tuned layers. The performance of the proposed models based on original and G-channel fundus images is discussed in the second experiment. The proposed method are compared with three related models in the last experiment. Ten-fold cross validation method is used to evaluate the performance of the proposed cataract classifiers.

4.2.1 The number of fine-tuned layers. In this section, the performance of the proposed method concerning the number of finetuned layers are investigated. In figure 3, the accuracy value of the proposed method with 1 (2, 3, and 4) fine-tuned layer (layers) is reported, which is labelled by "Fine-tuned 1 (2, 3, and 4)". In addition, the performance of fine-tuned AlexNet is also presented, in which only the last fully connected (FC) layer is trained. According to the results, it is easy to find that fine-tuning three layers in the proposed methods is a better choice for the classification of cataract. Moreover, the result of fine-tuned AlexNet demonstrates that only training the last FC layer cannot obtain better performance for fundus image classification.

4.2.2 Model construction on orginal and G-channel fundus images. In this experiment, we test the performance of original and G-channel fundus images for model construction. The results are listed in TABLE 2, which represent the classification performance

Table 2. : THE CLASSIFICATION RESULTS OF MODELS BASED ON ORIGINAL AND G-CHANNEL IMAGES

Feature set	se(%)	sp(%)	acc(%)
Original Images	90.14	92.39	92.07
G-channel images	93.42	95.73	95.37





Fig. 3: Performance of the proposed method concerning the number of finetuned layers.

of the proposed method on the two image sets, respectively. It can be found the model using green-channel images outperform the one based on original fundus images. The green-channel based model achieves the highest classification accuracy of 95.37% and it also obtains an improvement of 3.28% (3.34%) in se (sp). Therefore, it is concluded that the classifier on G-channel fundus images is more suitable for the cataract classification task.

4.2.3 *Model comparison.* In this section, the proposed method are compared with three related models. The classification performance of the four methods are shown in figure 4. Notably, the experimental data may not be different for the five methods. However, the comparison is meaningful for studying the performance of the proposed method. In figure 4, It can be found that the proposed method outperforms the three methods, which means that our model is also comparable and competitive with the three models.

5. CONCLUSION

In this paper, a novel deep learning based method is proposed to classify cataract using fundus images. To avoid relying on mass labelled data, deep transfer learning is used to reduce the number of parameters that need to be trained. In the proposed method, the first five layers of AlexNet are utilized for general feature learning and three subsequent layers are designed and fine-tuned for the classification of cataract. To obtain a desirable model, green-channel fundus images are applied to train the proposed method. Comparing with three related models, the proposed method achieves the highest accuracy of 95.37%, which is comparable and competitive. The experimental results reveal the proposed method is effective for the classification of fundus images.

In the future, we focus on deep learning methods with attentional mechanisms to build better models for cataract diagnosis.

6. REFERENCES

 Miguel Caixinha, Elena Velte, Mário Santos, and Jaime B Santos. New approach for objective cataract classification based on ultrasound techniques using multiclass svm classifiers. In Ultrasonics Symposium (IUS), 2014 IEEE International, pages 2402–2405. IEEE, 2014.



Fig. 4: Performance comparison with two related models.

- [2] Jia Deng, Wei Dong, R. Socher, Li Jia Li, Kai Li, and Fei Fei Li. Imagenet: A large-scale hierarchical image database. *Proc* of *IEEE Computer Vision & Pattern Recognition*, pages 248– 255, 2009.
- [3] Weiming Fan, Ruifang Shen, Qinyan Zhang, Ji-Jiang Yang, and Jianqiang Li. Principal component analysis based cataract grading and classification. In *E-health Networking, Application & Services (HealthCom), 2015 17th International Conference on*, pages 459–462. IEEE, 2015.
- [4] Allen Foster et al. Vision 2020: the cataract challenge. Community Eye Health, 13(34):17–19, 2000.
- [5] Xinting Gao, Huiqi Li, Joo Hwee Lim, and Tien Yin Wong. Computer-aided cataract detection using enhanced texture features on retro-illumination lens images. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 1565–1568. IEEE, 2011.
- [6] Xinting Gao, Stephen Lin, and Tien Yin Wong. Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, 62(11):2693–2701, 2015.
- [7] Liye Guo, Ji-Jiang Yang, Lihui Peng, Jianqiang Li, and Qingfeng Liang. A computer-aided healthcare system for cataract classification and grading based on fundus image analysis. *Computers in Industry*, 69:72–80, 2015.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Sun Jian. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision & Pattern Recognition*, 2016.
- [9] Moncef Khairallah, Rim Kahloun, Rupert Bourne, Hans Limburg, Seth R Flaxman, Jost B Jonas, Jill Keeffe, Janet Leasher, Kovin Naidoo, Konrad Pesudovs, et al. Number of people blind or visually impaired by cataract worldwide and in world regions, 1990 to 2010worldwide prevalence of cataract, 1990–2010. *Investigative ophthalmology & visual science*, 56(11):6762–6769, 2015.
- [10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

- [11] Huiqi Li, Liling Ko, Joo Hwee Lim, Jiang Liu, Damon Wing Kee Wong, and Tien Yin Wong. Image based diagnosis of cortical cataract. In *Engineering in Medicine and Bi*ology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, pages 3904–3907. IEEE, 2008.
- [12] Huiqi Li, Joo Hwee Lim, Jiang Liu, Paul Mitchell, Ava Grace Tan, Jie Jin Wang, and Tien Yin Wong. A computer-aided diagnosis system of nuclear cataract. *IEEE Transactions on Biomedical Engineering*, 57(7):1690–1698, 2010.
- [13] Huiqi Li, Joo Hwee Lim, Jiang Liu, Damon Wing Kee Wong, Ngan Meng Tan, Shijian Lu, Zhuo Zhang, and Tien Yin Wong. An automatic diagnosis system of nuclear cataract using slit-lamp images. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 3693–3696. IEEE, 2009.
- [14] Yitao Liang, Lianlian He, Chao Fan, Feng Wang, and Wei Li. Preprocessing study of retinal image based on component extraction. In *IT in Medicine and Education, 2008. ITME 2008. IEEE International Symposium on*, pages 670–672. IEEE, 2008.
- [15] Hualei Shen, Hongwei Hao, Lihong Wei, and Zhibin Wang. An image based classification method for cataract. In *Computer Science and Computational Technology*, 2008. ISC-SCT'08. International Symposium on, volume 1, pages 583–586. IEEE, 2008.
- [16] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
- [17] Ruchir Srivastava, Xinting Gao, Fengshou Yin, Damon WK Wong, Jiang Liu, Carol Y Cheung, and Tien Yin Wong. Automatic nuclear cataract grading using image gradients. *Journal* of Medical Imaging, 1(1):014502–014502, 2014.
- [18] Li Xiong, Huiqi Li, and Liang Xu. An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis. *Journal of Healthcare Engineering*, 2017, 2017.
- [19] Ji-Jiang Yang, Jianqiang Li, Ruifang Shen, Yang Zeng, Jian He, Jing Bi, Yong Li, Qinyan Zhang, Lihui Peng, and Qing Wang. Exploiting ensemble learning for automatic cataract detection and grading. *Computer methods and programs in biomedicine*, 124:45–57, 2016.
- [20] Meimei Yang, Ji-Jiang Yang, Qinyan Zhang, Yu Niu, and Jianqiang Li. Classification of retinal image for automatic cataract detection. In *e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on*, pages 674–679. IEEE, 2013.
- [21] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.
- [22] Jin Zheng, Liye Guo, Lihui Peng, Jianqiang Li, Jijiang Yang, and Qingfeng Liang. Fundus image based cataract classification. In *Imaging Systems and Techniques (IST), 2014 IEEE International Conference on*, pages 90–94. IEEE, 2014.