A diagnostic method of high myopia based on transfer learning

Baibing Chen* and Cheng Wan

College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Corresponding author: 1909307773@qq.com

Abstract. In order to promote the development of computer-aided diagnosis, improve the efficiency of high myopia diagnosis, and help healthcare workers assist in diagnosis, this paper uses 6571 fundus color images of high myopia and 6212 normal color images from the Geriatric Hospital of Nanjing Medical University as the dataset. The initialization parameters pre-training on ImageNet are loaded into the VGG-16, VGG-19, ResNet-50, ResNet-101, Inception-V3, and EfficientNet-B0 models, Learning rate changes dynamically according to the number of training epochs, and saving the model with the highest area under curve (AUC) value during the training process. Taking the best classification model from different networks, this paper achieves classification accuracy of 93.86%, sensitivity of 0.9126, specificity of 0.9677 and AUC value of 0.9866 on the test set. Each image takes 0.039 seconds, which is within the acceptable range meeting real-time of medically assisted diagnosis. Results show the fine-tuned ResNet-101 network meets the needs of medical assist diagnosis and can be used for screening high myopia, with high accuracy and real-time performance.

1. Introduction

Myopia is the visual health problem with the highest incidence in the world, the longest age span of patients, and the widest range of involvement [1]. In recent years, the age of myopia tends to be younger and the progress is accelerated, which make myopia become the main cause of children's visual impairment. At the same time, the incidence of high myopia and even pathological myopia also increases. Cataract, glaucoma, retinal detachment and myopic macular degeneration may lead to irreversible blindness, which can not be ignored [2]. According to the data of the National Health Commission, the overall myopia rate of children and adolescents in 2018 was 53.6%; myopia rate of 6-year-old children is 14.5%, myopia rate of primary school students is 36.0%, myopia rate of junior high school students is 71.6%; myopia rate of senior high school students is 81.0%; Among the myopia students in senior three, high myopia account for 21.9%. The pathogenesis of high myopia is complex. It is generally believed that environmental factors [3] and genetic factors [4] jointly participate in the occurrence and development of high myopia, and genetic factors play a very important role. Patients with myopia can develop into high myopia if they use their eyes incorrectly, and manual interventions can be used to slow down the progression of high myopia, as well as conservative or surgical treatments such as laser corneal refractive surgery and posterior scleral reinforcement if it is needed [5]. In recent years, deep learning develops rapidly. Convolutional neural network (CNN) have shined in the field of image classification and segmentation by virtue of their powerful feature extraction capabilities, and new CNN, such as ResNet and Inception, have improved and innovated on the original network to streamline model parameters and improve model accuracy. CNN are also widely used in medical image,
such as the grading of diabetic retinopathy [6], the diagnosis of glaucoma as well as cataract [7-8], the
detection of focal points [9] and the evaluation of fundus image quality [10], which has obtained results
of high accuracy, high specificity and high sensitivity.

The number of people with high myopia is large, while there are very few diagnostic algorithms for
high myopia. In order to reduce doctors' workload and improve the efficiency of diagnosis, this paper
proposes a diagnostic method of high myopia based on transfer learning, which can screen for high
myopia fundus diseases and help doctors to assist in diagnosis.

The algorithm flow of this paper is as follows:
(1) Transforming different types of images collected from different machines into the same type.
(2) Data set is randomly divided into 3:1:1 according to training set, validation set, and test set.
(3) The image is preprocessed and the redundant information is removed by clipping. In order to speed
up training, the size of the image is scaled to 512 × 512. In order to prevent over fitting, the image is
enhanced by data processing.
(4) Using VGG-16 [11], VGG-19 [11], ResNet-50 [12], ResNet-101 [12], Inception-V3 [13], and
EfficientNet-B0 [14] network to extract image features. Then training CNN by transfer learning. Finally
six classification models are obtained by fine-tuning CNN.
(5) Testing and evaluating the model on the test set, and selecting the best classification model.

2. Data and method
The overview of diagnosis of high myopia is shown in Fig. 1. Firstly, the fundus image goes through the
data pre-processing module before training. Next, we train models of different convolution neural
networks and save the model with the highest AUC value on the validation set. Finally, we select the
model with the largest AUC value from all the models and apply it to the screening of highly myopic
fundus images.

![Figure 1. Overview of the diagnostic algorithm for high myopia.](image)

2.1. Source of data
The fundus image used in this study is from the Geriatric Hospital of Nanjing Medical University. Due
to the machine change, the size and shape of the fundus images will be different, which contain image in
PNG with a resolution of 2592 × 1944, image in JPG with a resolution of 2544 × 1696 and image in JPG
with a resolution of 2196 × 1958. The data set was labeled by professional ophthalmologists of Geriatric
Hospital of Nanjing Medical University, including 6571 high myopia fundus images and 6212 normal
fundus images.

2.2. Data division
In this paper, data set is randomly divided into training set, validation set and test set according to the
approximate ratio of 3:1:1. The final data set divided are shown in Table 1.
Table 1. Data division.

|               | Training set | Validation set | Test set | Totals |
|---------------|--------------|----------------|----------|--------|
| High myopia   | 3918         | 1350           | 1303     | 6571   |
| Normal images | 3750         | 1207           | 1255     | 6212   |
| Totals        | 7668         | 2557           | 2558     | 12783  |

2.3. Data preprocessing

The image pre-processing mainly consists of data normalization and data enhancement. There are three different sizes of images in the dataset, and there are black borders which are invalid redundant information, so the black borders are removed first. Considering the size of the dataset and the computational limitations of the computer, we scale the image to 512×512 after the black border is removed, and the pre-processing process is shown in Fig. 2. Next, the image is flipped randomly horizontally and vertically with a probability of 0.5. In order to make the training process smoother, speed up the gradient descent for the optimal solution and even improve the accuracy, the input pixel values need to be standardized. As shown in (1) and (2), all three channel pixel values of the input data are standardized from [0,255] to [-1,1].

\[
\text{pixel} = \frac{\text{pixel}}{255} \quad (1)
\]

\[
\text{input} = \frac{(\text{pixel} - 0.5)}{0.5} \quad (2)
\]

Figure 2. Image preprocessing: (a) Retinal fundus image; (b) After cropping and scaling.

2.4. Model training

In this paper, the network is trained and fine-tuned using transfer learning, the goal of which is to apply knowledge or patterns learned on a domain or task to a different but related domain or problem, to migrate the knowledge structure from the related domain, and to complete or improve the learning of the target domain or task. The transfer learning strategy used is to initialize the VGG-16, VGG-19, ResNet-50, ResNet-101, Inception-V3, and EfficientNet-B0 networks by loading parameters pre-training on the ImageNet data set, which greatly saves the time and effort of training network models. To make the network converge quickly, all networks in this paper use SGD optimization algorithm, the momentum is set to 0.9. The loss function is a cross-entropy loss function; the initial learning rate is set to 0.0005, the learning rate is reduced to 1/5 for every 20 epochs of training. Due to memory limitations of graphics card, every batch of samples is set to 8 images, iterating 959 times every epoch. Training 100 epochs and saving the model with the highest AUC value on the validation set.

2.5. Test and evaluation

This experiment uses classification accuracy, specificity, sensitivity, receiver operating characteristic (ROC) curve and AUC value as evaluation indexes, considering the actual computer-aided diagnosis application, this experiment also adds to testing time for each image to see if it can meet the real-time requirements of aided diagnosis. The classification accuracy formula is (3), TP which indicates the number of images predicted to be highly myopic and actual label is highly myopic; FP which indicates the number of images predicted to be highly myopic but actual label is normal; TN which indicates the
number of images predicted to be normal and actual label is normal; and FN which indicates the number of images predicted to be normal but actual label is highly myopic. When the sample label categories are unbalanced, the classification accuracy at this point is not as convincing as the specificity and sensitivity. The specificity formula is (4), and the sensitivity formula is (5). Drawing the ROC curve according to the specificity and sensitivity and calculating the area under the curve.

\[
\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)
\]

\[
\text{Specificity} = \frac{TN}{FP+TN} \quad (4)
\]

\[
\text{Sensitivity} = \frac{TP}{TP+FN} \quad (5)
\]

3. Results
The classification results of high myopia fundus based on different convolutional neural networks are shown in Table 2. Comparing the evaluation indexes of the six methods, the specificity and AUC values of ResNet-101 are the highest among all CNN, with a classification accuracy of 93.86%, a specificity of 0.9677, and a sensitivity of 0.9126, which can help doctors to make auxiliary diagnosis and have a certain guarantee in the accuracy.

| Method     | Accuracy | Specificity | Sensitivity | AUC     |
|------------|----------|-------------|-------------|---------|
| VGG-16     | 93.47%   | 0.9387      | 0.9311      | 0.9783  |
| VGG-19     | 93.00%   | 0.9403      | 0.9207      | 0.9823  |
| Inception-V3 | 87.80%  | 0.7904      | 0.9563      | 0.9586  |
| ResNet-50  | 94.02%   | 0.9394      | 0.9230      | 0.9861  |
| **ResNet-101** | **93.86%** | **0.9677** | **0.9126** | **0.9866** |
| EfficientNet-B0 | 93.35%   | 0.9453      | 0.9230      | 0.9815  |

The AUC value is 0.9866, and time of 0.039 seconds is spent for per testing image, which meets the real-time requirements of computer-aided diagnostic systems. Compared the ROC curves of different methods (Fig. 3), ResNet-101 has the best AUC value. AUC value generally ranges from 0.5 to 1, and it is positively correlated with the prediction performance of a classifier, so we select ResNet-101 as final method.

![ROC curve of different methods](image)

4. Summary
Transfer learning is an idea as well as a means to increase the speed of training model and mitigate over-fitting problems caused by small amounts of data. Essence of CNN is a feature extractor, and high-dimensional features fused from low-dimensional features can be extracted by through CNN, different deep neural networks perform differently in different tasks, and there is no universal optimal
algorithm, only through continuous experimental attempts can we find the most suitable algorithm for the actual task.

In this study, an approach based on transfer learning for classifying highly myopic fundus images is proposed, with a simple training process and excellent evaluation indexes, which can also be applied to other classification tasks, such as Shin HC et al [15] who use transfer learning to classify interstitial lung disease (ILD) and demonstrated the role of transfer learning. Lam et al [16] who classify diabetic retinopathy achieve the highest sensitivity of 95% and specificity of 96% on Kaggle dataset, using transfer learning of the GoogLeNet models. Zago et al [17] who classify image quality achieve AUC of 0.9998 on DRIMDB dataset by fine-tuning CNN model. In this paper, the classification method of high myopia fundus images is realized by transfer learning. Using the high myopia classification model in this paper can realize the screening of high myopia and reduce the burden of doctor as a computer-aided diagnostic tool.

Acknowledgement
This research was financially supported by Chinese Postdoctoral Science Foundation (2019M661832); Jiangsu Planned Projects for Postdoctoral Research Funds (2019K226); Jiangsu Province Advantageous Subject Construction Project.

References
[1] I. G. Morgan, K. Ohno-Matsui, and S. Saw, “Myopia,” The Lancet, vol. 379, no. 9827, pp. 1739-1748, 2012.
[2] Y. Ikuno, “Overview of the complications of high myopia,” Retina, vol. 37, no. 12, pp. 2347-2351, 2017.
[3] V. J. M. Verhoeven, G. H. S. Buitendijk, F. Rivadeneira, A. G. Uitterlinden, J. R. Vingerling, A. Hofman, and C. C. W. Klaver, “Education influences the role of genetics in myopia,” European Journal of Epidemiology, vol. 28, no. 12, pp. 973-980, 2013.
[4] R. Wojciechowski, “Nature and nurture: the complex genetics of myopia and refractive error,” Clinical Genetics, vol. 79, no. 4, pp. 301-320, 2011.
[5] M. Chen, J. Dai, R. Chu, and Y. Qian, “The efficacy and safety of modified Snyder–Thompson posterior scleral reinforcement in extensive high myopia of Chinese children,” Graefe’s Archive for Clinical and Experimental Ophthalmology, vol. 251, no. 11, pp. 2633-2638, 2013.
[6] D. J. Hemanth, O. Deperlioglu, and U. Kose, “An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network,” Neural Computing & Applications, vol. 32, no. 3, pp. 707-721, 2019.
[7] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudigar, J. H. Tan, and U. R. Acharya, “Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images,” Information sciences, vol. 441, pp. 41-49, 2018.
[8] X. Xu, L. Zhang, J. Li, Y. Guan and L. Zhang, “A hybrid global-local representation CNN model for automatic cataract grading,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 2, pp. 556-567, 2020.
[9] K. M. Bijay Dev, P. S. Jogi, S. Niyas, S. Vinayagamani, C. Kesavadas, and J. Rajan, “Automatic detection and localization of Focal Cortical Dysplasia lesions in MRI using fully convolutional neural network,” Biomedical Signal Processing and Control, vol. 52, pp. 218-225, 2019.
[10] A. Raj, N. A. Shah, A. K. Tiwari, and M. G. Martini, “Multivariate regression-based convolutional neural network model for fundus image quality assessment,” IEEE Access, vol. 8, pp. 57810-57821, 2020.
[11] K. Simonyan, and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” International Conference on Learning Representations, 2015.
[12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
[13] S. Christian, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826, 2016.

[14] M. Tan, and Q. V. Le, “EfficientNet: rethinking model scaling for convolutional neural networks,” *International Conference on Machine Learning*, 2019.

[15] H. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285-1298, 2016.

[16] C. Lam, D. Yi, M. Guo, and T. Lindsey, “Automated detection of diabetic retinopathy using deep learning,” *AMIA Summits on Translational Science Proceedings*, vol. 2017, pp. 147, 2018.

[17] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, “Retinal image quality assessment using deep learning,” *Computers in Biology and Medicine*, vol. 103, pp. 64-70, 2018.