Research on OCT Image Diagnosis Algorithm for Diabetes

Yunxiang Liu¹,a, Jian Chen¹,b
¹Shanghai Institute of Technology SIT Shanghai, China
a yxliu@sit.edu.cn, b 894794146@qq.com

Abstract. Non-proliferative diabetic retinopathy (NPDR) is an early stage of diabetic retinopathy. Effective testing can stop the disease from developing. In this paper, a 7-layer convolutional neural network after migration learning based on MATLAB R2017a, was used to perform image recognition on the full-eye OCT images of normal and NPDR at various stages. The full-eye OCT images for the testing set and training set were obtained from an open source database at duke university. A total of 120 images were obtained by optical coherence tomography. There were four types of images, including 30 OCT images of normal, NPDR type 1, NPDR type 2 and NPDR type 3. We labeled 120 full-eye OCT images numerically and divided them into a training set and a testing set, 60 each. The first 15 pictures of each category are used as the training set, and the last 15 pictures are used as the testing set. Four images were randomly selected from the 60 testing set images to determine which of the four categories the extracted full-eye OCT images were. Finally, the accuracy of the convolutional neural network after transfer learning reaches the expected effect.

1. Introduction
Diabetic retinopathy (NPDR) is a common complication of diabetes that is irreversible and difficult to treat completely. According to statistics, as shown in Figure 1, the number of patients with diabetes in the world has increased from 285 million in 2009 to 425 million in 2017 [1]. As of 2017, the total number of diabetes patients in my country has reached 5.7% of the country's total population[2,3]. It is particularly important to detect and monitor the condition of NPDR in time. Doctors can find this series of lesions based on OCT. According to its severity, it can be divided into type I, type II and typeIII[4-6].

Figure 1 Number of diabetic patients in China
Deep learning (DL) is a new type of AI machine learning technology that uses some machine learning technologies to solve real-world problems and simulate human decision-making by developing neural networks.

In the past few years, deep learning (DL) has been widely used in medicine, aiming to imitate the neuron layer in the human brain to process and extract information, so that computers can learn without explicit programming. This technology can also be used to detect diseases, including fundus images of retinal diseases, chest radiographs of tuberculosis, and skin images of malignant melanoma.

Recently, deep learning has been used to identify risk factors related to cardiovascular disease (for example: blood pressure, smoking, and body mass index) from retinal photos. Kermany et al. reported in Cell that AI can detect several diseases, including diabetic macular edema, choroidal neovascularization, drusen, and pediatric pneumonia, with the possibility of diagnosis.

Screening patients with retinal complications in the diabetic population is a very important public health strategy, which aims at early detection of diabetic patients with intraocular complications, and receiving early eye care services and treatment.

In the United States, after screening 20 million diabetic patients, doctors found that about 1 million patients had DME, but such large-scale screening would bring a heavy financial burden. Diabetes patients are usually asymptomatic, but should be monitored so that appropriate treatment (for example, laser or anti-VEGF treatment) can be given before vision loss occurs. However, it is not very realistic to screen all elderly over 50 years old to identify and monitor asymptomatic individuals with early signs of AMD (ie drusen) or patients with advanced symptoms of AMD (ie CNV), but for these asymptomatic individuals Treatment is very effective in maintaining vision.

Deep learning (DL) is a new type of AI machine learning technology that uses some machine learning technologies to solve real-world problems and simulate human decision-making by developing neural networks.

At present, deep learning technology using ConvNets has been greatly developed. The transfer learning used by Kermany et al. is a method of using ConvNets to build an AI system, which has been pre-trained using a large data set in the public domain. For example, transfer learning allows the use of knowledge gained during training to identify animals in images for use in identifying retinal diseases from optical coherence tomography (OCT) images. ConvNet is composed of multiple layers of neurons, with trainable weights, so it can learn features and patterns. Inspired by the biology of the visual cortex of the brain, each neuron in ConvNet is connected to a local area of the input to understand the specific characteristics of the image. In medical imaging, there have been many publicly available ConvNet models (VGGNet, ResNet, Inception V3 and DenseNet) so far.

Deep learning (DL) is a new type of AI machine learning technology that uses some machine learning technologies to solve real-world problems and simulate human decision-making by developing neural networks.

Kermany et al. stated that it is expected to use deep learning and transfer learning technology to diagnose the three main retinal diseases diagnosed based on OCT images: DME, CNV and drusen.

In this study, the author used Inception V3 to train a deep learning framework, including ~37000 CNV images, ~11000 DME images, ~9000 drusen images and ~51000 images of healthy people; the total number of images is 4686 people Obtained. Subsequently, 1000 images were verified: 250 CNV images, 250 DME images, 250 drusen and 250 normal images.

Kermany et al. developed an AI framework for detecting CNV, DME, and drusen from OCT images. Their AI framework uses the transfer learning method Inception V3 on the training set, and uses three different methods to repeat 100 iterations, as shown in Figure 1. This method can diagnose diseases with 90% accuracy.

Among all three models, the diagnostic performance of deep learning is more than 90% accurate in distinguishing CNV, DME, drusen and normal images, and the best result is obtained in the binary classifier model (accuracy> 98%) ( figure 1). Although the accuracy is slightly reduced, the limited model can achieve accuracy of more than 90% even if the training set is 100 times smaller than the full data set.
Compared with six human experts, the deep learning system found similar results when identifying individuals who needed urgent referral based on OCT images. To further verify the effectiveness of the deep learning algorithm, images from 5826 patients (2538 images of bacterial pneumonia and 1345 images of viral pneumonia) were carried out on a set of 5,232 training sessions of children’s chest X-rays (CXR). And 1349 health images), and 624 images (234 health images and 390 pneumonia images) from 624 patients, and they achieved an accuracy of 92.8%.

Before entering the clinic, in order to further verify the effectiveness, the results should be directly compared with the existing deep learning system to measure the relative advantages, limitations, performance, efficiency and ease of use. In addition, with regard to the use of this method in other applications, the researchers conducted an occlusion test. This method successfully determined the area in ConvNet, which is important for making a diagnosis.

However, this may not be easily applicable to diseases with variable abnormal areas or other imaging modalities (for example: CXR). In addition, consider the institution where this method is most applicable: Is it suitable for general population screening in primary health care settings or to help ophthalmologists perform diagnosis in tertiary medical institutions? Finally, more generally, future research may address medical imaging challenges, such as how or when machines and human referees are different, and design methods to evaluate and explain the sources of error between humans and machines in quantity and quality.

In conclusion, Kermany et al. have shown that the deep learning system has excellent performance in diagnosing DME, CNV and drusen based on OCT images and diagnosing pediatric pneumonia based on children's chest X-rays. They also emphasized the application of transfer learning in small data sets.

However, there are still unresolved problems in many fields of deep learning medical imaging analysis. Therefore, machine learning and the medical community must cooperate closely, not only to promote the development and verification of deep learning technologies, but also to strategically deploy the application of these technologies in patient care.

AI performance is evaluated in three models: multi-class comparison, finite model, and binary classifier. For multiple group comparisons, the author uses AI to distinguish CNV, DME, and drusen images from healthy individuals. For the limited model, only 1000 images (250 CNV, 250 DME, 250 drusen and 250 normal) are used, and the number of images in these training sets is much smaller than the original training data set (116000 images).

For the binary classifier, the author divided the OCT image into CNV and normal, DME and normal, and drusen and normal to test individual AI algorithms in each case.

In the early 1990s, D. Huang et al. of Massachusetts Institute of Technology first proposed the concept of optical coherence tomography (OCT) [7]. Optical coherence tomography (OCT) is a cutting-edge optical diagnostic technology that has recently developed rapidly and is generally recognized by relevant people at home and abroad. It transmits a low-coherence light beam to the detection site to accurately inspect the corresponding tissue [8 -10]. And because of other advantages of OCT, such as non-radiation, non-invasiveness, no need to use contrast agents, high safety, and high sensitivity, OCT has been welcomed by physicians in medical fields such as ENT examinations [11]. At present, there are many algorithms that can be used to identify OCT graphs, including convolutional neural network (CNN) algorithms, support vector machine (SVM) algorithms, region growing (RSG) algorithms, etc. [12-16]. This article will use Convolutional Neural Network (CNN) algorithm to train and test OCT images on CPU and GPU.

2. Convolutional neural network applied to OCT image recognition

2.1. Supervised learning

Supervised learning is simply a learning method that needs to classify training samples. By categorizing, integrating and comparing the training set, after multiple judgments and calculations, the neural network outputs an optimal model, and then the optimal model of the neural network is compared with the model
to be tested in the future, thus the test set sort. Convolutional neural network is a typical supervised learning [17,18].

Supervised learning includes regression and classification.

Regression: Regression is to fit a curve of (X, Y). Where Y is a real number, making the value function \( L(f, (X, Y)) \) is the smallest.

\[
L(f, (X, Y)) = ||f(X) - Y||
\]  (1)

Classification: X and Y are real numbers. Before classification, we should name each classifier a different label. The process of labeling each classifier is called supervised learning among them \( f_i(x) = p(Y = i | X) \)

\[
L(f, (X, Y)) = -\log f_Y(X)
\]  (2)

2.2. Convolution operation

Convolutional layer (Convolutional layer) is the core part of CNN [19]. The parameters in the convolutional layer are obtained through BP optimization, and several convolutional layers constitute the hidden layer. The convolutional layer in CNN effectively extracts the features of the tested picture through the convolution kernel, and generally the first few convolutional layers can only get some of the most basic features of the image, such as edge features. As the number of convolutional layers increases, some minutiae features of each input image will also be extracted into feature maps [20].

Figure 2 is the most basic convolutional neural network convolutional layer architecture.

![Convolution layer, Convolutional neural network](image)

3. OCT image set CPU and GPU training test results

The CPU training results are shown in Figure 3.

![CPU training result](image)

GPU training results are shown in Figure 4.
4. Conclusion

This article aims to learn and apply a set of deep learning algorithms, combined with some retinal medical images, through the classification, labeling, and image preprocessing of 120 whole-eye OCT images, using a seven-layer CNN, through transfer learning. The final four-layer structure is fine-tuned to realize the recognition of medical images and judge the accuracy of the recognition, and the final OCT image recognition accuracy reaches 97.66%.

Here are some areas that can be further optimized.

References

[1] Wei R., Application of machine learning algorithms in the medical field [J]. Medical and health equipment., 2019, 40(03): 93-97.
[2] Ye L, Pharmaceutical care of patients with diabetic nephropathy and fundus hemorrhage [J]. Primary Medicine Forum., 2019, 23(16): 2350-2352.
[3] Xu L, Fundus image quality control in screening of diabetic retinopathy [J]. Beijing Biomedical Engineering., 2019, 38(02): 166-170.
[4] Li Z, Data prediction service construction method based on supervised learning [J]. Computer Technology and Development., 2019, 29(09): 188-194.
[5] Qian L, Study on the correlation between the degree of non-proliferative diabetic retinopathy and retinal blood flow index [J]. Chinese General Practice., 2018, 40(03): 93-97. 16(11): 1784-1786+1803.
[6] Lai X, Research on Deep Learning Unsupervised Learning Algorithm. [J]. Fujian Computer., 2019, 34(09): 102-103.
[7] Yu L, Application of optical coherence tomography angiography in the diagnosis of ocular diseases [J]. Journal of Shanghai Jiaotong University (Medical Edition)., 2018, 38(07): 829-834.
[8] Wu X., Optical Coherence Tomography Image Recognition of Retinopathy Based on Deep Learning [J]. Journal of Guangdong University of Technology, 2019, 38(02): 166-170.
[9] Gu C, Research on Artificial Intelligence Deep Learning in Fundus Image Recognition of Idiopathic Polypoid Choroidal Vascular Disease (PCV) [J]. Chinese Journal of Laser Medicine., 2019, 27(02): 147.
[10] Li C, Application of deep learning in image recognition [J]. Journal of Nantong University (Natural Science Edition)., 2019, 17(01): 1-9.
[11] Wang Y, Application and development of deep learning in image recognition [J]. China's strategic emerging industries., 2017, 48: 115-116.
[12] Sunshine. Research on license plate recognition technology based on support vector machines and deep learning[D]. Chengdu University of Technology, 2019.
[13] Zhao Yanqin, Tong Chaodi, Zhang Heng. License plate character recognition based on LeNet-5 convolutional neural network [J]. Journal of Heilongjiang University of Science and Technology, 2019, 29(03): 382-386.
[14] Liu Q, Application of optical coherence tomography vascular imaging technology in diabetic retinopathy [J]. International Journal of Ophthalmology., 16(04):678-680.
[15] Hu L, Research on k-space theory, algorithm and application based on phase contrast OCT [J]. Journal of Guangdong University of Technology.2014, 29(05):185-189.
[16] Cho N H., IDF diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projects for 2045 [J]. Diabetes Research & Clinical Practice, 2018
[17] Boone M, High-definition optical coherence tomography: adapted algorithmic method for pattern analysis of inflammatory skin diseases: a pilot study [J]. Computer-Aided Diagnosis. International Society for Optics and Photonics, 2015, 305(4):283-297
[18] Ginneken B V, A novel gas turbine fault diagnosis method on transfer learning with CNN [J]. Measurement, 2019:435-453
[19] Wang J, Convolutional neural networks for expert recommendation in community question answering [J]. Science China (Information Sciences). 2017(04)
[20] Sandra L, Descriptions of issues and comments for predicting issue success in software projects [J]. The Journal of Systems & Software, 2020, 168.