A Study of Ocular Disease Prediction Based on Convolutional Neural Network Algorithm

Zimo Zhang *

Department of Mathematics, University of British Columbia, Vancouver, Canada

* Corresponding Author Email: zimo21@ubc.student.ca

Abstract. With the widespread use of electronic devices and the accelerated development of an aging society worldwide, more and more people are facing the threat of eye diseases, which are often irreversible once they occur. Therefore, it is important to detect and prevent these eye diseases as early as possible. Many people would like to be able to detect potential disease risks in a timely manner by performing self-examinations at home, especially by performing preliminary analysis of fundus images and disease prediction, and by having the fundus examined on a regular basis. It is in this context that this study was conducted with the aim of developing a disease prediction model based on fundus images using a Convolutional Neural Networks (CNN) algorithm. By analyzing a patient's fundus, the model is able to make a preliminary assessment of whether he or she is suffering from a retina-related disease. After extensive training and testing, the model achieves a prediction accuracy of 90.3% on the test dataset, which is highly reliable and can help patients get some diagnostic help before going to the hospital and help doctors make a diagnosis to a certain extent. In the future, with the further development of the technology and the continuous improvement of the dataset, this deep learning-based ocular disease prediction model is expected to play a greater role in clinical practice.

Keywords: Eye disease; CNN algorithm; Eye fundus.

1. Introduction

In modern spciaty, excessive eye use is accompanied by poor eye habits, and eye diseases have become one of the most common diseases of old age, with a high probability of blindness which includes, glaucoma, cataract, age related macular degeneration, hypertension, pathological myopia, and other diseases or abnormalities. Diabetic retinopathy is characterized by blockage of retinal blood vessels [1], glaucoma is characterized by blockage of intraocular drainage [2], cataract refers to clouding of the lens [3]. Age related macular Ddgeneration is classified into two categories, dry and wet, and is characterized by thinning of the macula in the center of the eye and growth of abnormal blood vessels at the macula, respectively [4]. Hypertension refers to the destruction of the retina by high blood pressure. Hypertensive ophthalmopathy is a condition in which high blood pressure disrupts normal blood flow to the retina [5]. In addition, pathological myopia is also characterized by macular lesions in the eye [6]. All of these eye diseases can be determined by observing the state of the eye, and ocular radiographs have been very helpful to physicians in making a diagnosis. This has inspired research into whether deep learning methods can be used to help doctors analyze eye fundus.

In this paper, the convolutional neural network (CNN) technique, especially the pre-trained VGG19 model, was used to achieve early prediction of a variety of ocular diseases through feature extraction and classification analysis of a large number of fundus images. CNN is a deep learning algorithm specifically designed to process image data, and it can automatically extract useful features from images and perform classification. In the study of this paper, fundus diseases are analyzed automatically by this method, which effectively improves the prediction accuracy of the diseases.

2. Method

Convolution is a kind of mathematical operation, for the functions f and g, the expression of convolution (f * g) for them is shown as bellow.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) \, d\tau \tag{1}$$

Moreover, convolution is also widely used in image processing, natural language processing etc., the basic idea is to slide the kernel in the input data to carry out multiplication operations in local regions. A convolutional neural network (CNN) consists of three types of layers: a convolutional layer that is used to input data and extract features, a pooling layer that retains important information and reduces the spatial dimensions of the feature map, and a fully connected (dense) layer that performs classification [7]. In this study, a CNN is used on a large number of images of eye fundus and mainly used the preprocessed model which called VGG19.

3. Results

As mentioned before, this study focuses on predicting what kind of eye disease a patient has through eye fundus.

3.1. Introduction to Dataset

The dataset used for the study was Ocular Disease Intelligent Recognition (ODIR) [8]. ODIR is a structured ophthalmic database comprising 5,000 patients' ages, colour fundus images from the left and right eyes, and diagnostic keywords provided by specialists. This dataset is intended to reflect a "real-life" set of patient information gathered by Shanggong Medical Technology Co., Ltd. from various hospitals and medical centres in China. In these institutes, fundus images are obtained using a variety of market cameras, including Canon, Zeiss, and Kowa, resulting in a variety of image resolutions. Annotations were labelled by skilled human readers under quality control management. They classify patients under eight labels, which include: Normal (N), Diabetes (D), Glaucoma (G), Cataract (C), Age related Macular Degeneration (A), Hypertension (H), Pathological Myopia (M), Other diseases/abnormalities (O).

3.2. Exploratory Data Analysis

Before matching each image with its diagnosis, an exploratory data analysis was made for a dataset with some basic patient information. It is found that approximately 53.6% of the patients were male and 46.4% were female. In the gender distribution of patients with “normal fundus,” approximately 43.7% of patients were female and 57.3% were male. The histogram of the number of patients with normal fundus, cataract, diabetes, glaucoma, hypertension, myopia, age issues, and other ocular radiologic conditions is shown in Fig. 1, which shown that the dataset had not balanced the number of each eye disease and may cause some bias.

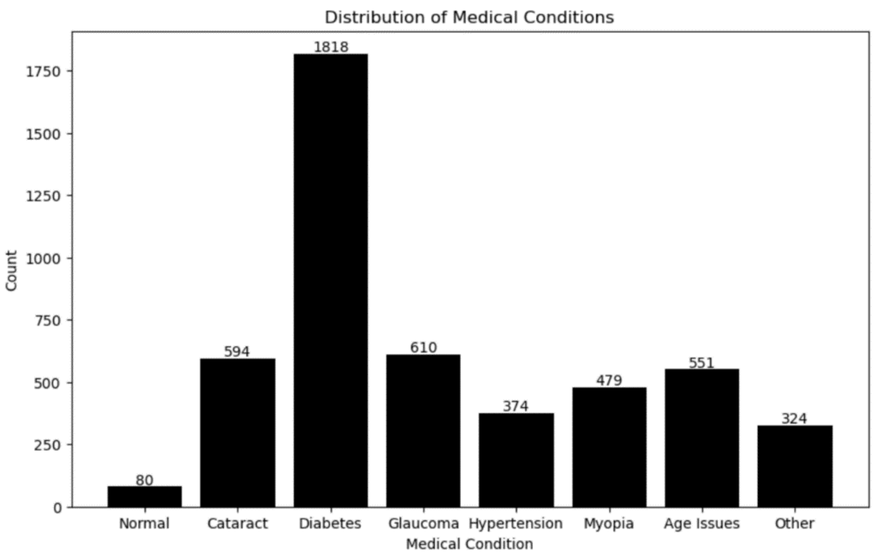

Fig. 1 Distribution of Medical Conditions

3.3. Dataset Preprocessing

In the data preprocessing stage, the concept of feature engineering was used to determine whether each row is characterized by a particular keyword, e.g., whether the diagnostic keyword for the patient's left eye in the first row is “cataract”. After determining each medical condition (including normal fundus, cataract, diabetes, glaucoma, hypertension, myopia, age issues, and other ocular radiologic conditions) in each row, each diagnosis was assigned to its corresponding fundus for model training and model testing. After these processes, the dataset was reduced to a combination of each ocular fundus image and its diagnostic classification, which can reduce the time and resource consumption for model training and prediction during the experiment. Feature engineering such as proposing disease signatures from diagnostic text selects the most relevant features to reduce dimensionality and improve model efficiency and performance. Finally, the dataset was splited into a training dataset and a test dataset.

3.4. Model Building

The deep convolutional neural network ---- Visual Geometry group (VGG19) was chosen as the model for this study. VGG19 consists of a 19-layer network [8] that reuses the same convolutional and pooling modules. The use of a small 3*3 convolutional kernel captures richer spatial hierarchical information. After multiple convolution and pooling layers, VGG uses several fully connected layers for final classification. The reasons for using VGG19 as the base model for the ocular disease recognition task include: VGG19 can learn a rich representation of image features due to its deep structure; VGG19 has been pre-trained on large-scale datasets, so it is highly generalizable. In our study, the pre-trained features can be used directly, and its excellent performance has been demonstrated to improve the accuracy and generalization ability of the model. The Sequential model is used to construct neural networks with linear stacked hierarchies.

In this study, after loading VGG19 as a feature extractor, all the layers of BVGG19 are frozen so that the weights of the pre-trained model are not updated, and then the Sequential model is used to stack VGG19 and the new fully connected layers to construct a deep learning model for a specific task, which makes the model more flexible and accurate.

3.5. Model Evaluation

Then the model was evaluated, in which the loss (the loss of the model on the test data, which hardens the difference between the model's predicted value and the wit) was 0.5654, and the accuracy (i.e., the percentage of the model's correct prediction) was 88.33%, which both reflect the model's high accuracy rate.

3.6. Prediction

Using the model trained in the method described above, the test data were predicted, and each ocular negative in the test data was analyzed to determine its medical status as shown in Fig. 2. Correct predictions were made in 90.3% of all predictions.

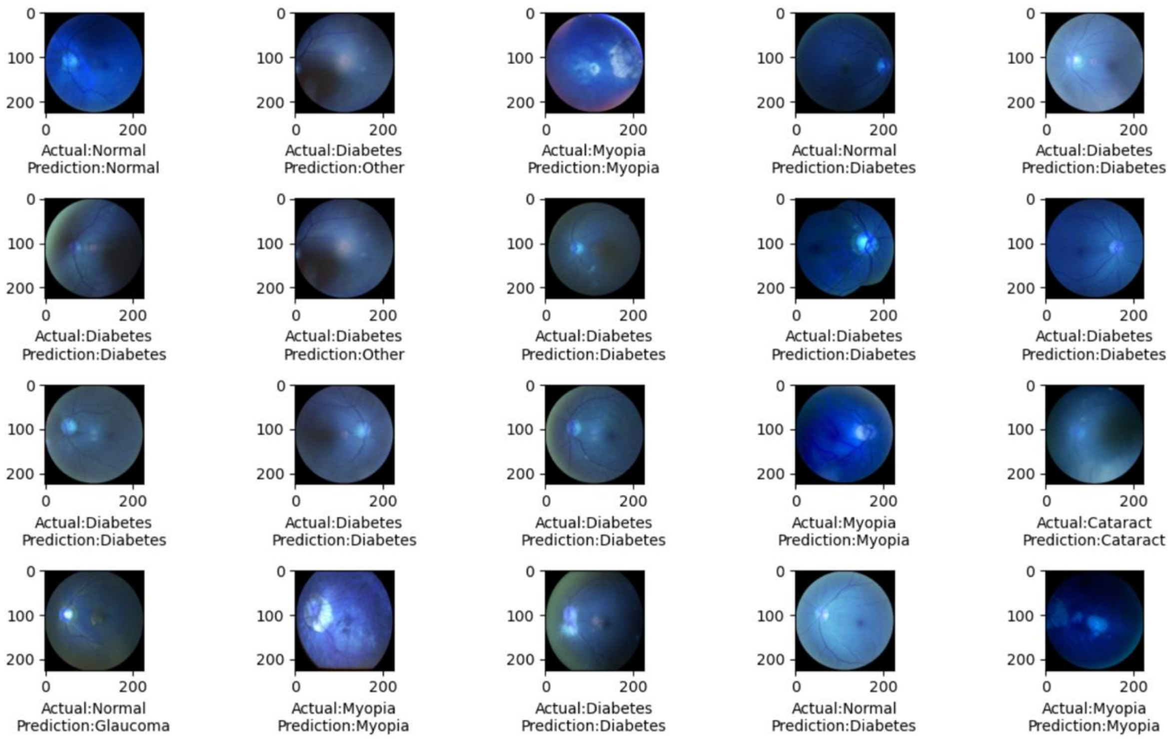

Fig. 2 visulization of the predicted diagnosis of each fundus

3.7. Discussion

The ocular disease diagnoses of the images in the predicted test dataset are shown in figure 2, and each diagnosis is assigned to its corresponding fundus. The previewed figure shows that most of the diagnoses are correct, except for “normal fundus” which can be easily misinterpreted as other diagnoses.

One of the major directions in current research on ocular diseases is to identify glaucoma patients by analyzing ocular negatives [9], and a KNN classifier was used to obtain an accuracy of up to 86%. However, in this study, only 200 pieces of data were used for model training, and the KNN algorithm used may lead to model overfitting due to its instance-based learning approach [10].

However, the 90.3% accuracy of this study is not sufficient for real-world healthcare applications for several reasons. First, in the dataset used for the study, the number of each medical condition type varies, with the highest number of 1818 data items and the lowest number of 50 data items, which may lead to the model being biased towards the majority of the classes. Moreover, the pre-training data of VGG19 comes from a generalized image database, which cannot cover all possible ocular disease manifestations. In addition, the features extracted by VGG19 may not be the best representation for ocular disease recognition, which can be solved by working with ophthalmologists to adjust the model. Regularization techniques such as early stopping can also be used to optimize the model and prevent overfitting.

Research has limited the performance of the CNN algorithm while controlling its computational cost. In future studies, researchers can ensure that the CNN algorithm is chosen in environments where research resources are not limited (e.g., deployment of GPUs) or use other simplified methods that have little impact on accuracy, such as parameter sharing, to control the cost. Once the model has been trained to a high level of accuracy, patients can use it in home to get the expletory diagnosis. Additionally, the model can also be integrated into the hospital healthcare system to assist doctors in diagnosing diseases and reducing the likelihood of medical errors.

In future studies, such as the glaucoma study mentioned above, the model selection and model training for a specific eye disease can be performed to ensure a prediction accuracy of 99% or more in that particular eye disease (due to the specificity of the medical industry) and apply it to the work of hospital doctors when ethical and moral conditions permit.

4. Conclusion

This study started from the need to predict a patient's eye health by learning the characteristics of diagnosed eye negatives. Due to the large amount of data, the resulting prediction accuracy was an unexpected 93.6%. Despite this high level of accuracy, the model is not yet ready for application. In future studies, modeling a specific eye disease and achieve a prediction accuracy of 99% or more for that particular eye disease can be a selectable path.

References

[1] Centers for Disease Control and Prevention. Common eye disorders and diseases. Centers for Disease Control and Prevention, 2020.

[2] Kalloniatis M, Bui B B, Phu J. Glaucoma: challenges and opportunities. Clinical and Experimental Optometry, 2024, 107(2): 107-109

[3] Liu Y C, Wilkins M, Kim T, et al. Cataracts. The Lancet, 2017, 390(10094): 600-612.

[4] Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the Age-Related Eye Disease Study Report Number 6. American journal of ophthalmology, 2001, 132(5): 668-681.

[5] Zimmerman R. You are here: Home/Cleveland Clinic Microvascular Complications of Diabetes Microvascular Complications of Diabetes.

[6] Morgan I G, Ohno-Matsui K, Saw S M. Myopia. The Lancet, 2012, 379(9827): 1739-1748.

[7] Sateesan A, Sinha S, KG S, et al. A survey of algorithmic and hardware optimization techniques for vision convolutional neural networks on FPGAs. Neural Processing Letters, 2021, 53(3): 2331-2377.

[8] Larxel. Ocular disease recognition. Kaggle, 2020.

[9] Bock R, Meier J, Michelson G, et al. Classifying glaucoma with image-based features from fundus photographs//Pattern Recognition: 29th DAGM Symposium, Heidelberg, Germany, September 12-14, 2007. Proceedings 29. Springer Berlin Heidelberg, 2007: 355-364.

[10] Li B, Yu S, Lu Q. An improved k-nearest neighbor algorithm for text categorization. arXiv preprint cs/0306099, 2003.