

A Deep Learning-Based Diabetic Retinopathy Classification and Diagnosis System

Shaojie Li *, Ying Guo

School of Information and Communication Engineering, Shenyang University of Technology, China

* Corresponding author: Shaojie Li (Email: 2375954588@qq.com)

Abstract: Diabetic Retinopathy (DR) was a chronic complication caused by diabetes. If not diagnosed and treated in time, it can severely affect vision and even lead to blindness. Due to the diversity and complexity of lesion areas, relying solely on manual detection make it difficult to establish quantitative judgment standards and introduces significant uncertainty. Therefore, developing an efficient automated diagnostic system was crucial. This paper proposed a diabetic retinopathy classification and diagnosis system based on the PyQt5 platform. The system was trained using the APTOS2019 dataset and employed a method combining Residual Networks (ResNet) with Focal Loss. This approach enabled the system to autonomously extract features from lesion areas and accurately classify the severity of the lesions. This not only helped to conserve medical resources and improve diagnostic efficiency but also provided reliable decision support for clinicians, thereby improving patient treatment outcomes.

Keywords: Diabetic Retinopathy; Residual Networks; Focal Loss; Classification System.

1. Introduction

Diabetic Retinopathy (DR) is an abnormal hemorrhage of the micro vessels in the fundus of the eye due to chronically high blood glucose levels in the body and dysfunction of insulin metabolism [1]. DR carries great medical and social significance as a serious complication faced by patients with diabetes mellitus. The incidence of diabetes has gradually increased with the changes in modern lifestyles, making diabetes and its related complications a prominent health problem worldwide. Although in the early stages of diabetes and DR, patients' vision may not be affected, as the disease progresses, especially to the advanced stages of DR, vision will suffer irreversible damage, so early retinal screening and effective diabetes management will be an important measure to alleviate the great health threat and social burden caused by diabetic retinopathy [2]. However, DR detection is a time-consuming and labor-intensive manual task, and with the rapid increase in the number of diabetic patients, the number of retinal images to be diagnosed in the healthcare system will also increase, which will put a huge burden on our healthcare delivery system. Based on the above discussion, this paper proposes a classification and detection system for diabetic retinopathy based on deep learning theory, which allows computers to independently find pathological features from fundus images and classify and diagnose the degree of lesions, thus being able to assist doctors in diagnosis and treatment.

2. Model Construction

2.1. ResNet Models

ResNet (residual network) is a deep convolutional neural network model proposed by He et al [3] in 2015. In traditional deep convolutional neural networks, the problems of gradient vanishing and gradient explosion occur as the number of layers increases, leading to training difficulties and bottlenecking of network performance. In order to solve this problem, ResNet proposes a residual learning idea, i.e., the output is added to the input to obtain the residual, and used as

the input for the next layer of the network. This can make the network easier to train, and can effectively alleviate a series of problems arising from network deepening. Specifically, the residual module contains two convolutional layers and a jump join, where the first convolutional layer is used to reduce the dimensionality of the feature map, and the second convolutional layer is used to restore the dimensionality of the feature map and add it to the input feature map to obtain the residual. Let the input be x . The computation process of the residual module is:

$$H(x) = F(x) + x \quad (1)$$

where $F(x)$ denotes the convolution operation in the residual block and $H(x)$ denotes the final output. Figure 1 shows a schematic structure of the residual unit.



Figure 1. Structure of residual unit

2.2. Focal Loss Function

Focal Loss [4] is a loss function designed for the category imbalance problem, proposed by Tsung et al. in 2017. It is mainly used to solve the situation where there are a large number of easy-to-categorize samples (DR images without lesions and in the non-proliferative stage) and a small number of difficult-to-categorize samples (DR images in the proliferative stage) in the DR classification task, and by introducing the focal factor and the balancing parameter, the model focuses more on the difficult-to-categorize samples, thus improving the model's performance in the category imbalance scenario.

The mathematical expression for the focal loss function is:

$$\text{FocalLoss}(p, y) = -\alpha(1 - p)^\gamma \log(p) \quad (2)$$

where p is the predicted probability of the model, y is the true label (0 or 1), α is the balance parameter, and γ is the focus factor. The focus factor is mainly used to adjust the contribution of easy-to-categorize samples to the loss, and the balance parameter is used to adjust the weights of positive and negative samples. When the samples are easy to classify, the focus factor makes the loss weight of the easy-to-categorize samples decrease, thus reducing their impact on model training; while for the samples that are difficult to classify, the smaller value of the focus factor makes the loss weight of the difficult-to-categorize samples increase, and pays more attention to this part of the samples. As shown in Figure 2 demonstrates the relationship between the predicted and true values in the loss function in neural networks.

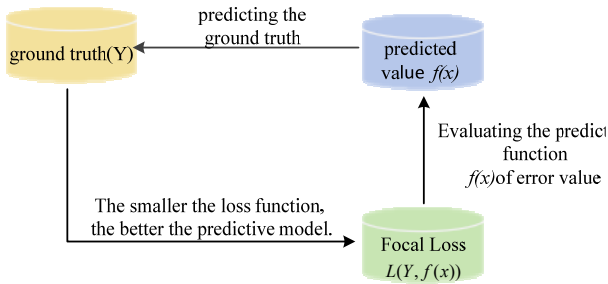


Figure 2. Relationship between predicted and true values in the loss function

3. Experimental Results and Analysis

3.1. Experimental Environment and Parameter Settings

The dataset used in this paper is the publicly available medical imaging dataset, APTOS 2019, provided by the “APTOS 2019 Blindness Detection” competition organized by Kaggle [5], which contains 3,662 fundus images covering five types of lesions ranging from normal to 5 lesions, including mild non-proliferative lesions, moderate non-proliferative lesions, severe non-proliferative lesions, and proliferative lesions. Meanwhile, for the problem of uneven image contrast and variable image quality in the dataset, this paper adopts the method of restricting the contrast adaptive histogram and the method of weighted fusion of the original image and the image after Gaussian filtering to enhance the image features and make the lesion areas more obvious to improve the overall image quality, and uses data enhancement means during the training process, including mirroring, rotating, translating, and brightness enhancement, etc. . Finally, in this paper, the preprocessed dataset is divided into training set, validation set and test set according to the ratio of 8:1:1. epoch is set to 50 times, learning rate is set to 0.001 and batchsize is set to 32.

The evaluation metrics in this paper are: accuracy (Accuracy, Acc), recall (Recall), specificity (Specificity, Spe), and consistency test (Kappa).

3.2. Results

Table 1. Comparison of results of different measures in the models

Evaluation metrics	LeNet	DenseNet	GoogleNet	Proposed
ACC	0.95078	0.9374	0.9523	0.9601
Recall	0.951	0.93996	0.9472	0.9534
Spe	0.951	0.98348	0.98788	0.9887
Kappa	0.93816	0.921031	0.93967	0.94358

Table 1 shows the comparison of the classification results obtained in this paper after using residual network (ResNet50) with focal loss function with the results of classical convolutional neural network (LeNet) [6], dense connectivity network (DenseNet) [7], and multi-scale convolutional network (GoogleNet) [8] for the five classifications of DR. As can be seen from the table, the model used in this paper has higher accuracy for the five classifications of DR.

4. DR Classification Detection System Design

By loading the completed and saved model weights in Section 3.2 that performed optimally in the validation set, this paper designs a DR classification interface that can directly display unknown samples.

4.1. Module Design

PyQt is a Python module for creating cross-platform desktop applications, a Python bundle for the Qt library. Developers can utilize Python to write efficient and aesthetically pleasing desktop applications using the GUI, multithreading, networking, and database features in Qt. However, in the process of designing a GUI interface, relying only on pure code to write the position and color of text and images in the interface can make the whole design process abstract and complicated, so both PyQt and PySide provide a unified tool Qt Designer to assist the design. Qt Designer is a visual design tool in the Qt suite, which can be used to design a GUI application by directly Qt Designer is a visual design tool in the Qt suite, which can be used to edit the interface by directly dragging and dropping components and setting properties. Afterwards, the generated .ui file can be converted to corresponding Python code by the pyuic tool, which makes the integration of design and code more convenient. The combination of the two provides developers with a comprehensive toolset that supports visual interface design and feature implementation, making it a powerful tool for creating modern desktop applications.

4.2. Design Flow of DR Classification Detection System

The specific design flow of the diabetic retinopathy detection system is shown in Figure 3.

First of all, the basic design of the graphical user interface, such as button controls, text controls, etc., is carried out by Qt Designer, a tool provided by PyQt; after that, the basic graphical interface that has already been designed is loaded by uic. loadui and some practical functions are added to the interface, such as selecting the lesion image to be detected from the computer's system disk, pre-processing the DR image, detecting the DR lesion grade, and saving the pre-processed image back to the computer, Detecting the DR lesion level, and saving the pre-processed image back to the computer, etc.; and then use the clicked. connect function to connect signals and slots, where signals are used to pass events between objects, and slots are functions that respond to these events. When an object sends out a signal, other objects can listen to this signal, if any object connects to the signal, their corresponding slot function will be executed, the two are the mechanism used for event processing in the Qt framework; finally, the detection of any DR lesion image can be realized through the designed classification detection interface.

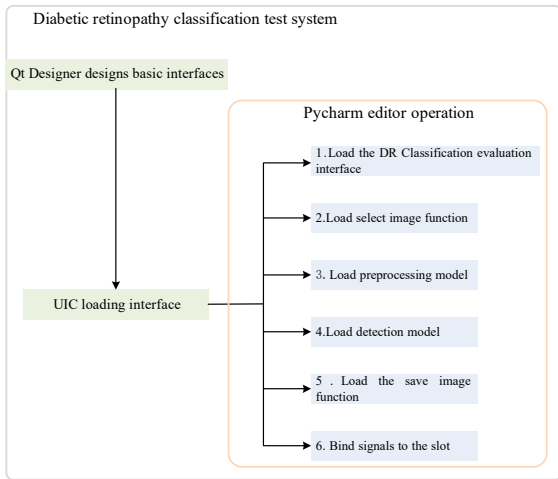


Figure 3. Design flow of diabetic retinopathy detection system

4.3. Implementation of DR Classification Detection System

(1) Initial interface design

In this paper, Qt Designer is used to design the initial outline of the system interface. Specifically, as shown in Figure 4. The left side of the interface is the control box, which is used to provide the basic controls used for GUI interface development, such as operation buttons, text input boxes, etc. The right side of the interface includes the object viewer for displaying the objects and their categories of the controls in the main window, and the attribute editor to allow the user to view and change the value of the object's attributes; the center of the interface is the main window, which is used for the design of the graphical interface and the components selected from the control box can be placed in the control box. The center of the interface is the main window, which is used for the design of the graphical interface and for placing the components selected in the control box. The interface designed for the DR classification test consists of fourteen QLabel controls for displaying the results, three Line controls for planning the interface, and four QPushButton controls.

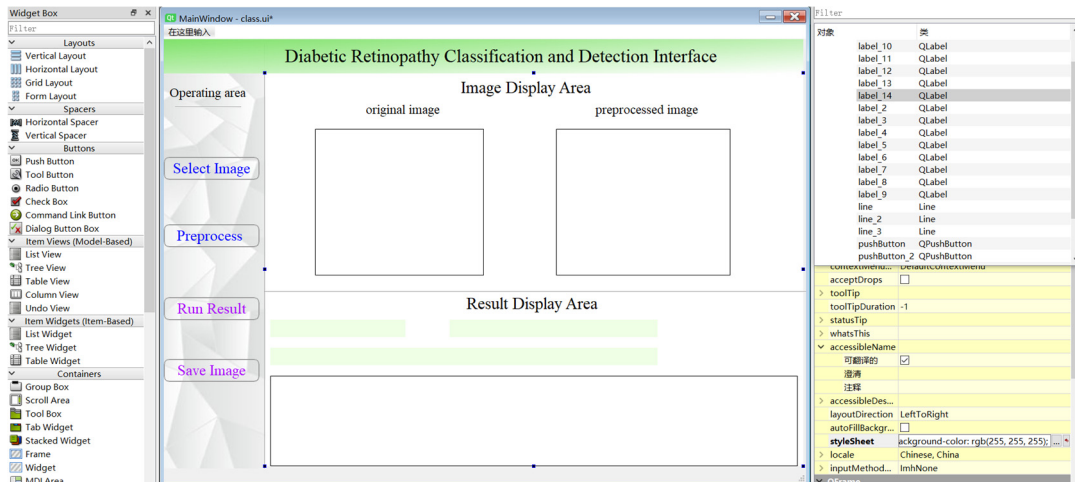


Figure 4. System lesion detection interface

(2) Obtaining the detection model

The process of result detection is shown in Figure 5: first load the classification labels DR0, DR1, DR2, DR3, DR4 of DR, and load the weights of the optimal model of performance, at this time, the loading of the model proposed in this paper has been completed, and then you can select a picture that you want to detect in the system, and finally click on the “Run Result” button in Figure 4 to enter the evaluation mode. After that, we can select a picture we want to detect in

the system, and then click the “Run Result” button in Fig. 4 to enter the evaluation mode, in which the Softmax function is loaded to obtain the probability value of the picture belonging to the five lesion stages, and then the Argmax function is used to obtain the index corresponding to the maximum probability, and then the corresponding label is found through the index, and then the classification and identification of the picture to be detected can be completed at last.

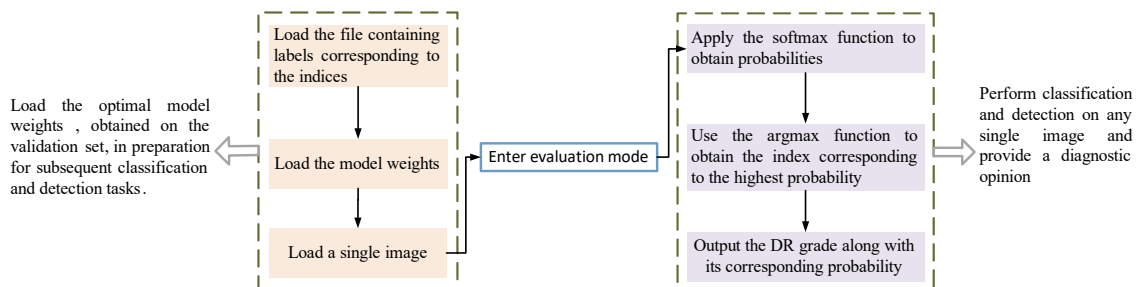


Figure 5. Detection process

Firstly, load the system lesion detection interface designed in Qt Designer through uic. loadui, then through QFileDialog.getOpenFileName to enable the operator to

select the lesion image to be detected on the computer, then load the preprocessing model and detection model to classify and diagnose the image, and through clicked. connect

function to bind the designed control to the above model, and finally use `QFileDialog.getSaveFileName` to save the detected image. connect function to bind the designed control to the above model, and finally use `QFileDialog.getSaveFileName` to save the detected image.

4.4. Interface Run Results

Figure 6 shows the classification results for an unknown sample entered. In Figure 6, click on the “Select Image” button to enter the computer system disk to select the image to be detected; click on the “Preprocess Fundus Image” button to process; click on the “Run Result” button to get the DR category label, detection result, probability and diagnostic opinion. Click on the “Run Result” button to get the DR category label, detection result, probability size and diagnostic opinion.

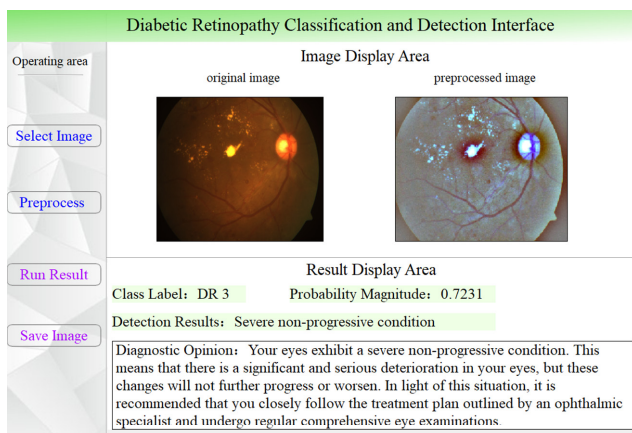


Figure 6. System lesion detection run results

5. Conclusion

In this study, a classification detection system for diabetic retinopathy based on residual network and focus loss function

is proposed. The system adopts a deep learning method to learn and extract features from a huge dataset, which avoids the influence of human subjective factors and thus ensures the accuracy of diagnostic results. In addition, the system has the ability of automation, which can significantly improve the efficiency of diagnosis and save doctors' time and energy, thus reducing medical costs and improving the efficiency and quality of medical services.

References

- [1] ZHANG Huirong, XIA Yingjie. Analysis of factors related to visual prognosis in patients with retinal vein occlusion[J]. Chinese Journal of Ophthalmology, 2002, 4(02): 37-41.
- [2] YAN Chunyan, YUAN Yingmei, GUO Aimei et al. Literature analysis of the current status of research on delayed medical care for diabetic patients in China[J]. Geriatrics Research, 2023, 4(5): 38-41.
- [3] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [4] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.
- [5] Aravind Eye Hospital, APTOS 2019 blindness detection [EB/OL]. <https://www.kaggle.com/c/aptos2019-blindness-detection>. 2019-6-29/2023-4-24.
- [6] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [7] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. Computer ence, 2014, 9(4): 1-8.