OXFORD BROOKES UNIVERSITY

Bella Ma

Multi-Class Diabetic Retinopathy Classification using Self-Attention Residual Networks

Oxford Brookes University in collaboration with Chengdu University of Technology
Supervised by Dr Happy Nkanta Monday

Abstract

Diabetes is one of the most common metabolic diseases worldwide, accompanied by mild to severe secondary complications, including diabetic retinopathy (DR), which can damage the retina and lead to vision loss. Image processing systems for the screening of diabetic retinopathy (DR) have been developed to partially address the increasing screening demand as the number of diabetic patients globally continues to rise. However, the accuracy of diabetic retinopathy diagnosis remains a key issue. This project aims to improve diagnostic accuracy by developing a new convolutional neural network (CNN) model based on residual learning combined with self-attention mechanisms, utilizing deep learning technology.

Dataset

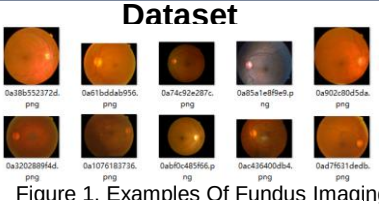

Figure 1. Examples Of Fundus Imaging

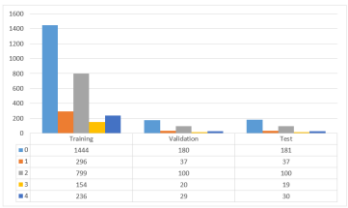

Figure 2. Image Partition Scale

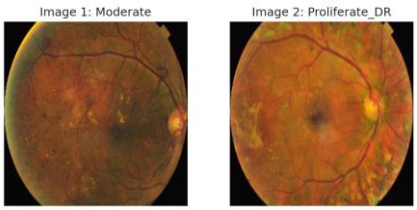

Figure 3. Enhancement with CLAHE

In the process of image preprocessing, the training set is first randomly scaled and cropped, and the CLAHE (Contrast Limited Adaptive Histogram Equalization) enhancement technique is used to process the images of diabetic retinopathy.

Ensemble models

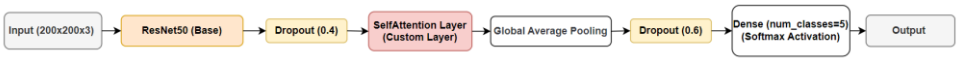

Figure 4. Ensemble model

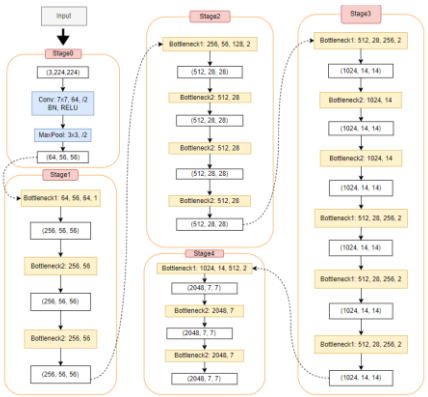

Figure 5. ResNet-50 Network Structure

This model integrates deep convolution, residual blocks and self-attention mechanisms, providing a robust solution for the complex task of image classification of diabetic retinopathy. This model takes the pre-trained ResNet50 as the basic model, which is helpful for extracting advanced features from the input images. ResNet50 has been pre-trained on a large amount of data and is capable of capturing complex image features.The customized self-attention layer enables the model to focus on the most important parts of the input data. This mechanism helps the model better understand the key features of the data, thereby improving the classification performance. This model contains multiple Dropout layers, which helps to reduce overfitting problems during the training process. By randomly eliminating some neurons, the model is forced to learn a more robust feature representation.

Model evaluation

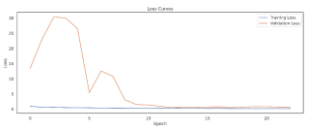

Figure 6. The loss of the model

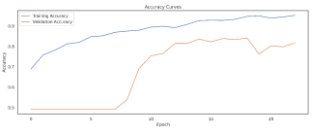

Figure 7. The accuracy of the model

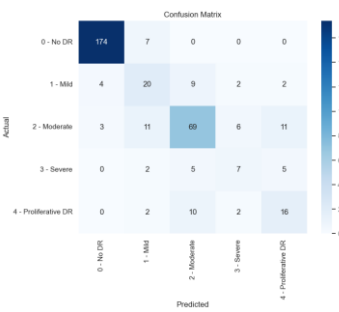

Figure 8. The confusion Matrix of the model

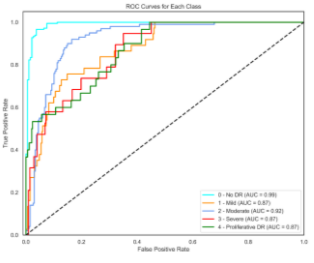

Figure 9. The ROC Curves of the model

The above figure shows the results of the prediction and classification of diabetic retinopathy using this model, presenting each result of the model.

Interpretability and significance analysis of the model

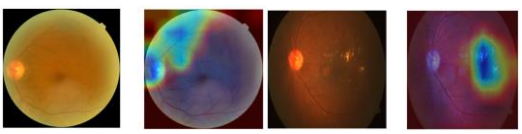

Figure 10. The heatmap of level '0 - No DR'

Figure 11. The heatmap of level '1 - Mild'

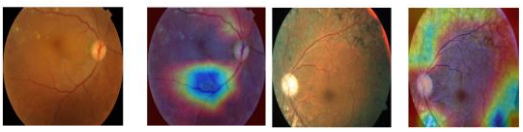

Figure 12. The heatmap of level '2– Moderate'

Figure 13. The heatmap of level '4 - Proliferative DR'

Grad-CAM generates class activation maps by calculating the gradient information of image categories [25], revealing which regions in the image are the most important for the model's prediction. Generally speaking, brighter colors (for example, red, yellow) indicate that the model pays more attention to the area, while darker colors (for example, blue, purple) indicate less attention.

WEB APPLICATION DEVELOPMENT

Figure14. Web interface to the model

Figure 15. Result page

The deployment runs in the browser, but only provides the web front-end interface and the flask runtime environment, while the inference prediction still runs in the local terminal.

Future work

1. More advanced integration methods and attention mechanisms will be explored in the future to further improve model performance in complex image classification tasks.
2. For post-deployment model optimization, strategies such as model lightweighting, incremental learning and model monitoring can be used.