DIABETIC RETINOPATHY CLASSIFICATION WITH TRANSFER LEARNING

A PREPRINT

Banupriya Valluvan
Department of Information Technology
University of Stuttgart
Stuttgart
st175151@stud.uni-stuttgart.de

Hanjie Li
Department of Electrical Engineering
University Stuttgart
Stuttgart
st146744@stud.uni-stuttgart.de

February 7, 2022

ABSTRACT

Diabetic retinopathy is an eye illness in which the retina is damaged as a result of diabetes. Since it is critical to have a routine eye checkup with the doctor which turned out to be expensive and time consuming, this project attempts to develop a mechanism for predicting diabetic retinopathy[1]. The Indian Diabetic Retinopathy Image Library (IDRID) comprises of retinal fundus images that was acquired from an Eye Clinic located in India. The goal is to determine if a patient has non-referable (NRDR) or referable diabetic retinopathy (RDR) based on the photographs (requiring immediate specialist attention). This project uses ResNet50 that has been pre-trained. Additional experimental findings contributed to an overall test accuracy of roughly 84%.

1 Introduction (by Banu)

Diabetic retinopathy(DR) is the most prevalent cause of visual loss in diabetics. A challenging difficulty is that the (IDRID) dataset has a limited number of images, making it difficult to achieve good performance. The goal of the project is to classifiy the images into non-referable diabetic retinopathy (NDRDR) and referable diabetic retinopathy (RDR) according to their features.

2 Model (by Hanjie and Banu)

Since the IDRID image dataset has small scale data, the concept of transfer learning is considered to be used in this project to improve the training efficiency and avoid overfitting. The pre-trained model ResNet50 provided by tensorflow keras is deployed as a part of the model architecture. In addition, an average pooling layer and a fully connected layer is combined with the pretrained model ResNet50. The weights of the ResNet50 is obtained from ImageNet training and are set as non-trainable parameters. The parameters in the connected layers are trainable and the weights are updated in the training loop. The softmax activation function and categorical crossentropy loss function are used. More details about the training are shown in chapter 3.3 Training loop. The model architecture deployed is shown in Fig 1.

3 Experiment (by Hanjie)

3.1 Dataset

The Indian Diabetic Retinpathy Image Dataset (IDRID) is a dataset of retina fundus images. The part of the database "B. Disease Grading" is used within the scope of the project. The IDRID provides information on the disease severity of diabetic retinopathy and diabetic macular edema for each image[2]. The images are firstly classified into two classes according to the severity of diabetic retinopathy: images with labels 0 and 1 (0: having no, 1:mild non-proliferative DR) are classified into non-referable diabetic retinopathy(NRDR, 0 as integer label) and images with labels 2 and

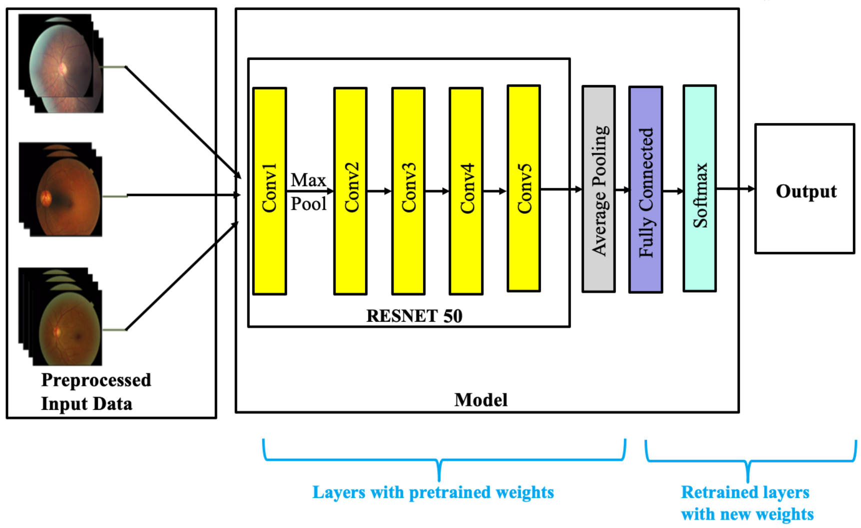

Figure 1: The customized ResNet50 deployed in diabetic retinopathy classification

up (2: moderate, 3: severe, 4: proliferative DR) are classified into referabele diabetic retinopathy (RDR, 1 as integer label). Since the dataset is highly imbalanced, resampling was necessay to be applied to the original dataset. The class distribution before the balancing is shown in Fig 2. In addition, 20% of the training data is splitted into a validation dataset. There are 300 images (150 images for each class), 83 images and 103 images in the train dataset, validation dataset and test dataset, respectively.

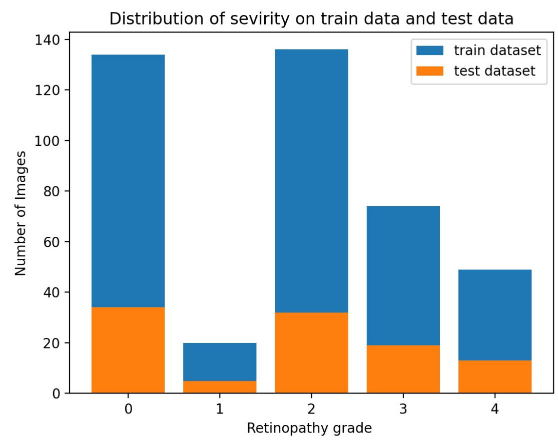

Figure 2: Label distribution based on the severity of DR in original IDRID dataset

3.2 Input pipeline (by Hanjie)

Images Preprocessing The original input images have different brightness. In some area the image is totally dark which contains almost no information. To make the relevant area easier to be spotted on, the uninformative area of each input was cropped by using center-cropping and specified bounding-box cropping.

Data Augmentation Due to the size of IDRID dataset, overfitting was observed in the training progress. To increase the diversity of the training set, data augmenation is applied to the train dataset. The following random transformations are applied to the IDRID train dataset:

- Random horizontal and vertical flipping

- Random cropping
- Random rotation
- Random brightness, contrast, hue, saturation

3.3 Training Progress

3.3.1 Training Loop (by Hanjie)

The train dataset is already repeated infinitely in the input pipeline. The total training steps, log intervals and checkpoint intervals are configurable in the gin file. For each batch, an operation recorder GradientTape scope is created.

Inside the GradientTape scope The model is called and the loss is computed. The loss between the labels and predictions is computed by the categorical cross-entropy function.

Outside the GradientTape scope The gradients of weights of the model with regard to the loss are retrieved. The Adam optimization is used as the stochastic gradient descent method in the custom training. The learning rate is set to 0.0001. In addition, the train confusion matrix and validation confusion matrix are updated and can be monitored with Tensorboard.

3.3.2 Saving Model (by Banu)

To save the model, checkpoints are written periodically into the local disk. Compared to the SaveModel format, checkpoints capture the accurate values used by the model without any description of the computation defined by the model. The checkpoints can be restored for the continuous training and evaluation.

3.4 Evaluation

Metrics (by Banu) **Confusion matrix** is selected as a metric because the confusion matrix not only reveals the errors that classifier is making, but also the types of errors that are being made. This breakdown assists in overcoming the drawbacks of relying on categorization accuracy.

Deep Visualisation (by Hanjie) To make the deep neural network more intuitive and understandable, deep visualisation is used in the project. Class activation mapping(CAM) is one of most commonly used methods to interpret the prediction decision made by CNN and highlight the importance of image region to the prediction[3]. In this project Gradient-weighted Class Activation Mapping (Grad-CAM) and Guided Grad-CAM methods are deployed. The results of deep visualisation are shown in Fig 3.

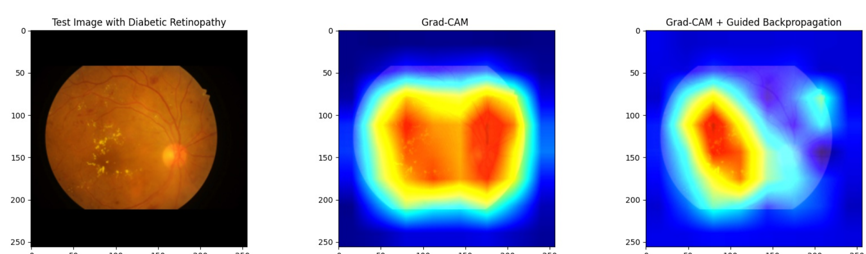

Figure 3: Deep visualisation with Grad-CAM and Guided Grad-CAM

Profiling (by Banu) Profiling aids in understanding the hardware resource consumption (time and memory) of the different TensorFlow operations (ops) in the model, allowing to identify performance bottlenecks.

4 Results (by Hanjie and Banu)

The evaluation results on test dataset are shown in fig 4. The best test accuracy achieve is 84.0% . The confusion matrix for the binary classification is shown in fig 5. Using pretrained ResNet50 and data augmentation, the overfitting problem can be reduced and the accuracy can be improved.

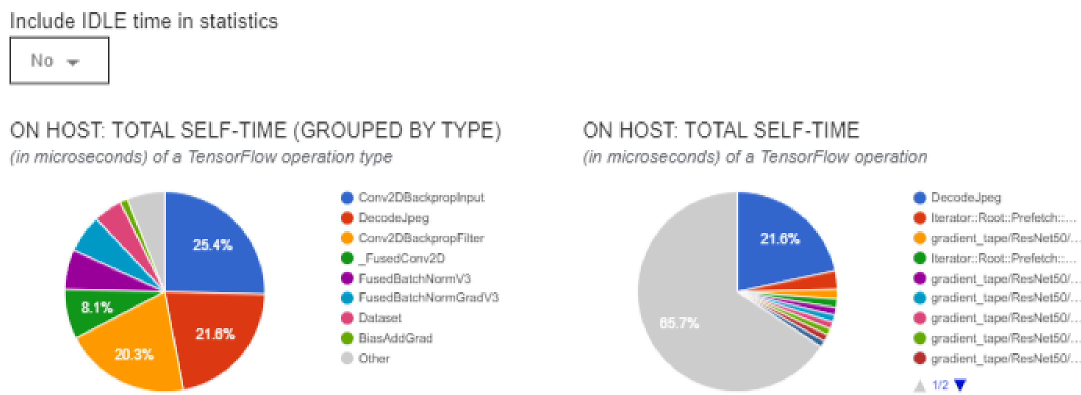

Figure 4: Profiling

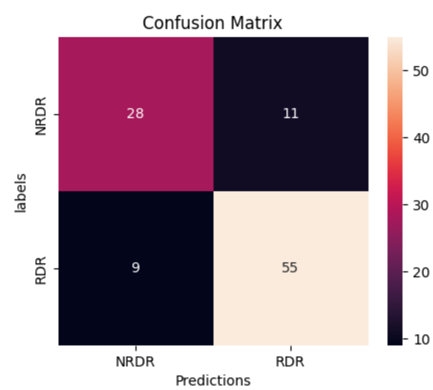

Figure 5: Confusion matrix

References

[1] Chua Kuang Chua Choo Min Lim EYK Ng Muthu Rama Krishnan Mookiah, U Rajendra Acharya and Augustinus Laude. Fast classification of handwritten on-line arabic characters. In *Computers in biology and medicine*, page 43(12):2136–2155, 2013.

[2] Prasanna Porwal, Samiksha Pachade, Manesh Kokare, Girish Deshmukh, Jaemin Son, Woong Bae, Lihong Liu, Jianzong Wang, Xinhui Liu, Liangxin Gao, et al. Idrid: Diabetic retinopathy–segmentation and grading challenge. *Medical image analysis*, 59:101561, 2020.

[3] Ramprasaath R Selvaraju, Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, and Dhruv Batra. Grad-cam: Why did you say that? *arXiv preprint arXiv:1611.07450*, 2016.