A NETWORK DEPENDING ON RESNET FOR DIABETIC RETINOPATHY CLASSIFICATION

Universität Stuttgart

ZHONG Liangyu, CAO Shijia

Institute of Signal Processing and System Theory, University of Stuttgart, Germany

1. Introduction

Diabetic retinopathy (DR) is an illness which causes damages to retina.
The goal of this project is to build a generic framework to classify non-referable (NRDR) and referable (RDR) diabetic retinopathy based on fundus images with deep learning algorithms.

2. Object and Methods

2.1 Dataset

We train our network on the Indian Diabetic Retinopathy Image Dataset (IDRID), in which severity grade of diabetic retinopathy of retinal fundus image is provided.

2.2 Input pipeline

- Preprocessing
  - Redefine the labels for binary classification
  - Use the left and right border of retina to crop the images and resize them to 256×256
  - Upsampling to balance the dataset
- Use TFRecord to load data efficiently
  - Serializing the preprocessed data
  - Reducing the disk space usage 212.2MB ⇒ 21.9MB
- Data augmentation
  - Method: Rotation, Shift, Zoom, Brightness, Flipping, Shearing
  - To avoid overfitting

2.3 Model

ResNet is a kind of convolutional neural network with skip-connection, which has shown high accuracy for classification tasks. Our network is a simplified version of ResNetV2 [1].

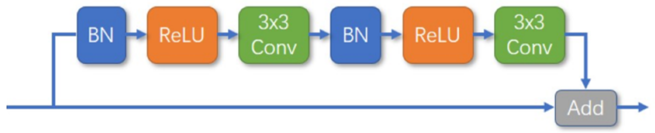


Abbildung 1: *A ResBlock in ResNetV2*

7×7 Conv(strides=2), BN, ReLU
3×3 Max pooling(strides=2)
ResBlock(strides=2) ch=16
ResBlock(strides=2) ch=32
ResBlock(strides=2) ch=64
ReLU, Global average pooling
Dense(unit=2), Softmax

Tabelle 1: *Model architecture*

We use sparse cross-entropy as our loss function.

2.4 Metrics

- Confusion matrix
- Accuracy

2.5 Code optimizing

- Logging with abseil
- Configuration with gin
- Accelerate training with autograph, average epoch time: 11.92s→7.18s.

2.6 Training and evaluation

- Use Adam optimizer to train our network for 300 epochs
- Save checkpoint for every 5 epochs
- Hyperparameter(HP) tuning
  - Grid search
  - Hyperparameters:
    - learning rate: 1e-2, 1e-3, 1e4 and 1e-5
    - batch size: 2 and 4

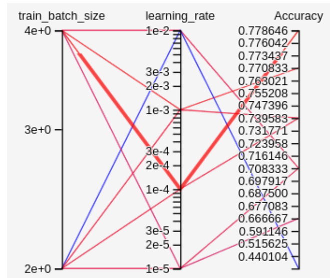


Abbildung 2: *HP tuning*

- Based on above best hyperparameters, if it's still underfitting as figure 3, use small learning rate=1e-5 and stronger data augmentation for further training(50 epochs).

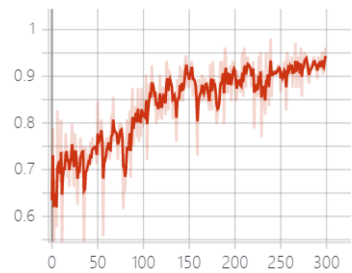
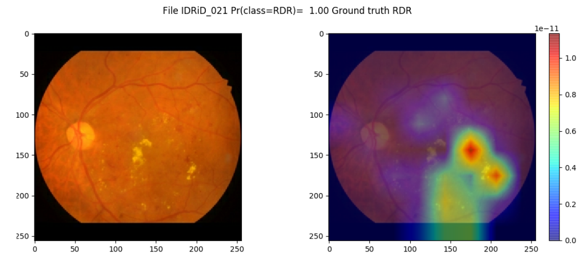


Abbildung 3: *Validation Accuracy*

2.7 Deep Visualization

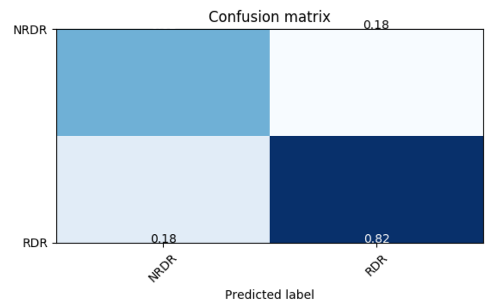
- Method: Grad-CAM
- The derivatives of prediction with respect to convolution output show network's attention.



E.g., for RDR images, our network mainly focuses on hard exudates on retinal fundus, which is reasonable.

3. Result

Overall, the network which we build can complete the DR classification task with a good performance.
Final accuracy is 82.2% and the corresponding confusion matrix is shown as follows.



Literatur

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. *CoRR*, abs/1603.05027, 2016.