DIABETIC RETINOPATHY DETECTATION, TEAM 2

Xu Siying, Li Liangzhu

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

1. Input pipeline

- Build datasets
  - Read data: Images are first read by file names for later preprocessing. Labels are read and saved in one-hot coding form.
  - Shuffle data index: Since the array of filenames has lower dimension than the array of images, we shuffle the index of filenames, thereby reducing the running time of the preprocessing stage.
  - Parse function: Firstly, it will transform a pair of input data (filename, label) to the corresponding data pair (image, label). Secondly, images will be decoded, resized and normalized and then transformed by one data augmentation method based on the random number it receives.
  - Datasets: including datasets of images and labels. The dataset is split into 383 samples for training and 30 samples for validation. The test dataset contains 103 samples and will not be shuffled.
- Online data augmentation: The data augmentation function is called during training to obtain infinite random data.

2. Model Architecture

- Self defined model
  - Model architecture: The model has totally 19 layers, including in total 5 convolutional layers, 5 max pooling layers, 5 dropout layers, 3 dense layers and 1 flatten layer.
  - Performance: To prevent overfitting, the training process is early stopped when training iteration reaches 900 times. The test accuracy can reach 0.76.
- Transfer learning
  - Model architecture: Based on the Xception model, two dense layers are added after removing its original last optional dense layer.
  - Layers: The Xception model has a total of 133 layers. The first 100 layers are set as untrainable and the next 33 layers are set as fine-tuning layers.
  - Performance: After about 150 training steps, the training accuracy can reach 0.99, the validation accuracy is about 0.9, and the test accuracy can reach 0.825. The performance has been significantly improved compared to the self defined model.

3. Training and Evaluation

- Training routine
  - Self defined model: we have written our own training routine to train the self defined model instead of using fit method, the optimizer is Adam with learning rate of 0.001. The training iteration steps are set to be 1600, but the real iteration numbers will be decided by the validation accuracy and loss, in order to prevent overfitting.
  - Transfer learning model: use fit method to train the transfer learning model. The optimizer is also Adam. The learning rate is 1/20 of self defined model.
- Evaluation
  - Deep visualization:
    - Method: Gradient-weighted Class Activation Mapping, which can highlight the important regions in an image for a specific prediction.
    - Performance:

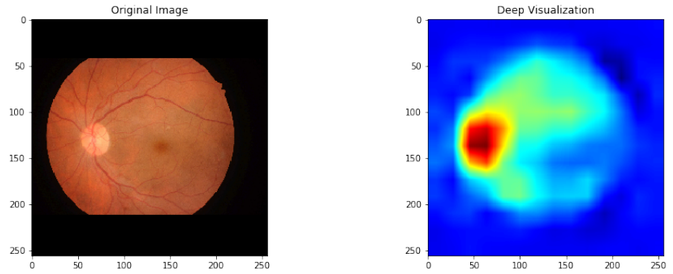

Abbildung 1: *Deep visualization by Grad-CAM*

  - Tensorboard: By using tensorboard, we track and visualize accuracy and loss. We also use Profiler in tensorboard to analyze the performance of input pipeline.
  - Checkpoint: check points are saved every 10 steps, which captures all parameters of the model.
  - Confusion matrix: through the confusion matrix, we can intuitively compare the correctness of binary classification of different models and methods.

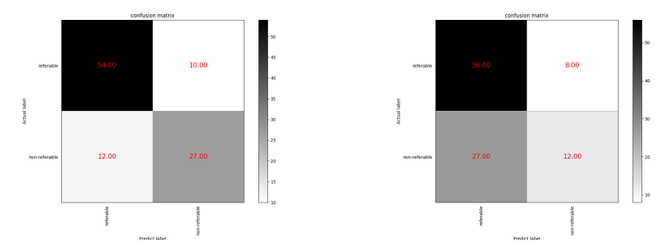

Abbildung 2: *Confusion matrix (a) with data augmentation (b) without data augmentation*

Deep Learning Lab 2020, February 4, Stuttgart, Germany