MSBD6000B Project 2 MAK, KIT HO 01652721

1. Data Preprocessing

We downsize the images to 128x128, and normalize each intensity to the range [0, 1] (by dividing every number by 255). Before feeding each image to the neural network in each epoch, we add random rotations (max 30 degrees), random scaling (width/height with a factor from 0.2 to 1) and random flipping (up-down and/or left-right) to the image.

2. The Model

Our model is based on ResNet18, which is the smallest of the famous ResNet family of models [1].

The ResNet is mainly built using two kinds of blocks, namely the convolutional blocks and the identity blocks (we used the non-bottleneck version). The identity block is constructed by first applying two convolutional layers (with stride=1 for the 2nd stage and stride=2 for later stages) with batch normalization and relu activation, and then add back to the input. The convolutional block is similar, except we apply a convolution for the input before adding.

The ResNet18 consists of 5 stages. The first stage is a standard convolutional layer with 64 7x7 filters of stride 2, followed by a 3x3 max pooling of stride 2. This is followed by 4 stages of (convolutional block -> identity blocks) layers repeated 2 times in each stage. The filter size of the blocks in each stage is 64, 128, 256, 512 respectively. We then apply a 2x2 average pooling for the output of stage 5 for feature extraction purpose.

On top of the ResNet18, we added a fully connected layer of 512 nodes, with dropout probability of 0.5 before connecting to the output (dense layer representing the 5 classes, with softmax). You may also see a sketch of the architecture in the file “model.png” in the same repository.

We train the model from scratch using our training set, for 100 epoches with early

stopping. The validation accuracy stalled at around epoch 70. We then stopped the training process. The training takes more than 10 hours with a single CPU core, and the validation accuracy is about 80%.

3.Reference

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, *Deep Residual Learning for Image Recognition* (2015). Available at https://arxiv.org/pdf/1512.03385.pdf