DGM Project Report 2

Helge Meier, Xiaoyan Xue, Yubao Ma, Zhiqian Yu
June 25, 2024

1 What we have worked on this week

- We have analyzed the file contents of MedMNIST, including the datasets breastmnist, bloodmnist, etc. Each dataset has 'train', 'val' and 'test', and the image formats of different datasets are also different, including grayscale images and RGB images. However, most of them are RGB images, so we focus more on 3-channel datasets.

- For the training part, we modified the main, model, and train parts of the original function and added a data.py file to load the MedMNIST data and compute its tensors as the features of the input neural network.

- During the training process, we used the method of small batch training, which not only helps to save memory space, but also allows the best model to be updated at any time. We set 3 points to save the existing model, 1, checkpoint: For example, save a checkpoint every 1000 epochs of training to prevent the program from losing progress due to unexpected interruptions. 2, latest: Ensure that there is always a latest state to restore, it always saves the most recent training state. 3, best: In order to ensure that the final model is the best performing model on the validation set, it is only updated when the performance on the validation set reaches a new high. Through the above methods, we can get the instant model and the best model so far.

- During training, we use classification training methods, such as training grayscale images and RGB images separately, to ensure that the input size of the neural network is uniform each time.

2 Results, findings and problems

- Training for RGB images works fine, but it still doesn't work for grayscale images. Due to the different number of input channels, we cannot build a neural network with a unified structure to be compatible with both sizes of input. A possible solution is to unify the input size, such as converting a single-channel to a three-channel, but this will increase the load on the system memory.

- Unfortunately we cannot train all the data together because we occasionally run out of memory. The solution is to train in batches. We have two GPUs, NVIDIA 4090 with power 400W and 4060 ti with 125W, and computational load of 0.4s/batch and 1.5s/batch respectively. For the entire training, it takes 5 6hours/50,000epochs and 25h/50,000epochs respectively. It can be seen that the time cost is huge, so this is also a very tricky problem.

3 Future work

1. All data including grayscale images and RGB images can be trained normally.

2. Fine-tune the neural network to get the best model.

3. Follow the task distributions:
   a) The grayscale images should be able to be trained by 02.07.2024 (1 week).
   b) Adjust parameters to optimize the model till 09.07.2024(1 week).