Thai and English OCR- Assignment1 (LT2326)

In this repository there are three scripts. The first one called data.py and generates the data needed for training and evaluating the model. The format of resulted the data is a tuple (language, dpi, style, path to image) in a txt file. To run the script, use the command below:

python3 data.py --lang Thai English --dpi 200 300 --style bold --tr_ratio 0.8 --te_ratio 0.1 --val_ratio 0.1 --input_path path to/ThaiOCR/ThaiOCR-TrainingSet --output_path path to output folder

The second script is called train.py and inside the script the images are processed accordingly (convert them to tensors, resized them to 128x128) and after the model is trained. To run the script use:

python3 train.py --tr_set /path to /training_set.txt --val_set /path to/validation_set.txt --batch 32 --epochs 10 --save_dir /path to directory for saving the model --log_file / path to/logs.log

The third script is called eval.py and evaluates the model and calculates the precision, recall, F1 score and the overall accuracy. To run the script use:

python3 eval.py --test_set /path to/testing_set.txt --model_path /path to/model.pth --batch 32

There is also another script called model.py which contains the model I trained. The model is a simple CNN architecture with one convolutional layer which extracts the features. It also consists of one fully connected layer for classifying the extracted features into different labels.

Experiments

All experiments were run on my local computer and not on mltgpu (I suppose that there will not be any problem running the commands). The data were split into training 80%, testing 10% and validation 10% sets. The model trained with 10 epochs, a batch size of 32 and a learning rate of 0.001. There is a possibility for those proportions, number of epochs and batch size to change through the command line.

1. Thai normal 200 dpi – Thai normal 200 dpi
   For this experiment the precision was 95%, recall: 92%, F1: 93% and the overall accuracy: 92.21%

2. Thai normal 400 dpi – Thai normal 200 dpi
   For this experiment the precision was 94%, recall: 93%, F1: 93% and the overall accuracy: 93.47%

3. Thai normal 400 dpi – Thai bold 200 dpi
   For this experiment the precision was 93%, recall: 92%, F1: 92% and the overall accuracy: 91.93%

4. Thai bold all dpi – Thai normal all dpi
   For this experiment the precision was 92%, recall: 91%, F1: 91% and the overall accuracy: 90.74%

   The challenging part of this assignment for me was to figure out how to handle the data in order to split them and later on to use then for training the model.