

MSc Individual Project Progress Report

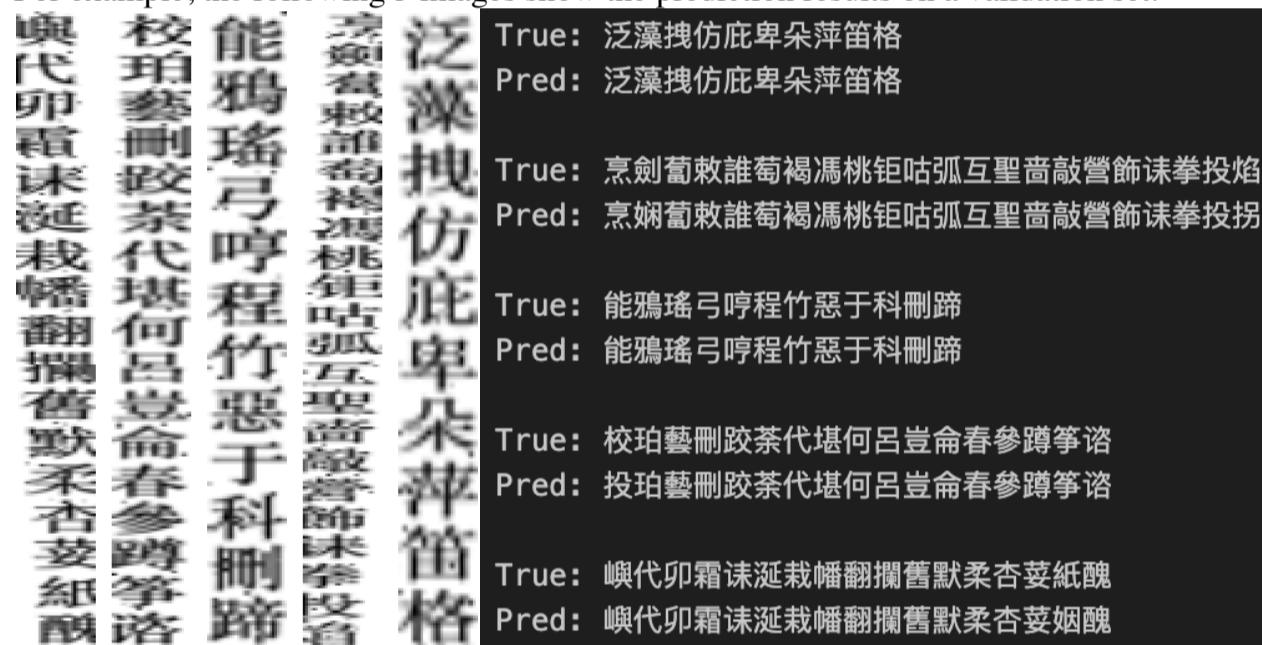
Michael Song (02405765) on 18 June

Progress before last meeting (24 May)

- Background and progress report
- Data collection (journal images with labels)
- Synthetic datasets for training
- Trained following models:
 - text line detector (LD)
 - character detector (CD)
 - character recognizer (CR)
 - character classifier (CC)

Progress after 24 May

- Compared the performance with following methods (however, none of them seems promising):
 - CR
 - CD + CC
 - LD + CR
 - LD + CD + CC
- Experimented with a CTC-based text recognition (TR) method (CRNN), that process the output from LD and outputs text content directly.
 - For example, the following 5 images show the prediction results on a validation set:



- The current solution to the task is LD + TR, as it is expected to give the best result
 - For example, the LD performs well on the journal images:



- The next step is to use the TR to predict the content in the text lines, in a sensible order
 - The evaluation is not finished yet

Future work

- Finish evaluation on the journal images, prevent over-fitting
- Experiment with other TR methods, such as encoder-decoder-based models