Proposal Topic

OCR in Student Identity Cards

Group 8
Tran Duc Nam 20215228
Tran Le Phuong Thao 20200604
Trinh Giang Nam 20215229
Nguyen Chinh Minh 20215224

1 Introduction

OCR stands for Optical Character Recognition. It allows users to scan text, numbers, and images and convert them into digital data. This technology is especially useful for identity card recognition, as it can be used to accurately detect and store someone's identity information with minimal effort. OCR can be used for HUST Student ID card recognition by using a scanner to scan the ID number on an ID card.

2 Objectives

- Understanding multiple deep learning models used in both 2 phases (Scene Text Detection Text Recognition) such as Convolutional Neural Network (CNN) and Transformer.
- Implementing some techniques in computer vision such as Non-Maximum Suppression (NMS), and image registration.
- Training available neural network for our own dataset and inferring output images for evaluation.
- Creating graphic user interface (GUI) for importing input images and generating text.

3 Expected Results

Our problem is divided into 2 sub-problems: Text Detection & Text Recognition for Vietnamese text in input image. The output is the coordinate of bounding boxes containing text and generated text label.

Figure 1: Input & Output of each phase.

4 Methodology

To information extraction from Student ID Card, we conduct with 2 phases: Scene Text Detection & Text Recognition.

- **Scene Text Detection.** Recently, segmentation-based methods are quite popular in scene text detection, as the segmentation results can more accurately describe scene text. However, the post-processing of binarization is essential for segmentation-based detection, which converts probability maps produced by segmentation method into bounding boxes of texts. We will implement a module named Differentiable Binarization (DB), which can perform the binarization process in a segmentation network.
- **Text Recognition.** We train VietOCR model combining CNN and Transformer. VietOCR has a property of generalization, even effectively performance in a new dataset by only using pre-trained weights. After extracting the coordinates of bounding boxes of relative text phases in image input, we crop the section containing text for using as an input of text recognition.

5 Experiments

5.1 Dataset

We use a dataset from Vietnamese Scene Text Detection And Recognition Challenge, which held up in 2021 by BK.AI Center. There are 3 datasets as follows:

- Training data: 2500 images and labels (consisting of 500 images provided by BKAI and 2000 images provided by VinAI)
- Validation data: 235 images and labels

- Testing data: 300 images and labels

The inputs for our model are raw images. Label of each image is store in file .txt. Each line of file containing labels have the coordinates of bounding box and text of its. It follows this format: x1, y1, x2, y2, x3, y3, x4, y4, label.
Our proposal problem is OCR in HUST Student ID Card. We will collect about 200 images of student cards and manually label them.

5.2 Results

The performance metric for evaluating results is CER (Character Error Rate). CER represents the percentage of characters in the text phase of ground truth compared with the predicted outcome which has an error. The lower CER, the higher performance of the model.

$$CER = \frac{S + D + I}{N}$$

where S: is the number of error characters compared with ground truth, D: the number of characters losing in prediction compared with ground truth, I: is the number of characters needing to add to the predicted result, N: is the number of characters in the ground truth.

6 Conclusion

OCR on Student Identity Cards is highly applicable in HUST's autonomous operation. For preparing data, we collect student ID cards and manually label them for training model. During our capstone project, we will understand all modules in deep learning architectures for building text detection and text recognition models. Furthermore, we utilize some important techniques in computer vision for improving the performance results.