NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

School of Computer Science and Engineering
College of Engineering

Real-time Scene Text Recognition

Overview

A program which employs a text detection and text recognition model in a pipeline to recognize scene text in real-time through a live camera feed. The program utilizes a pre-trained EAST text detection model and a self-trained CRNN text recognition model, of which the latter was the main focus of the project.

Technical Details

- Pytorch
- OpenCV2

CRNN model's accuracy performs at 87.75% against the ICDAR13 dataset and 78.10% against the IIIT 5K-words dataset. Program runs at an average of 20 FPS.

Ordered Displaying of Recognized Texts


```
==========================================
| Detected Texts by Groups               |
==========================================
| 0 | please take                        |
| 1 | nothing but                        |
| 2 | pictures                           |
| 3 | leave nothing                      |
| 4 | gbut foot prints                   |
==========================================
```

The program is also able to arrange the detected texts in the order in which they are meant to be read from the image. If the image contained multiple contexts of texts, the program is able to segregate them accordingly as well and not treat all of the words in the image as a single paragraph.

Student: Lim Eng Keat Joshen

Supervisor: Asst. Prof Lu Shijian