Helping the Visually Impaired Navigate at Bus Stops

Tang, Yu Han Brandon

Introduction

Outlined in the Land Transport Authority's "Land Transport Master Plan 2040" is the concept of "transport for all". The Visually Impaired (VI) form a sizeable subset of commuters who face considerable difficulties in taking public transport. These difficulties stem from infrastructural barriers like the lack of audio cues at bus stops and social barriers such as the VI being afraid of asking for help from sighted commuters.

In this project, we develop a proof of concept for an all-in-one system to enable the VI to independently gather information on the identity of buses that arrive at bus stops. It takes a video feed, detects the bus number of buses that come to the bus stop and reports this via audio.

Our implementation uses objection detection[1], optical character recognition[2] (OCR) and speech synthesis algorithms[3] from Microsoft Azure.

Methodology

Workflow

Understanding the Azure Services

- Read documentation
- Experiment with proof of concept scripts

Preparing Models

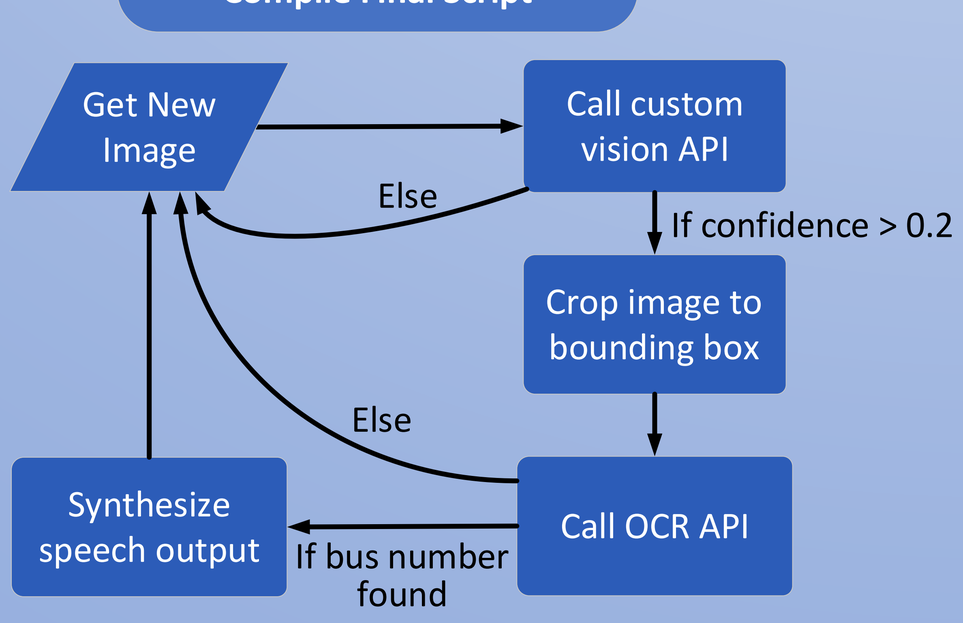
- Video arrival of buses
- Splice videos into individual frames
- Annotate bounding box of bus numbers & destinations
- **Train Custom Vision Algorithm**

Compare OCR Algorithms
1. Read API
2. OCR API
3. Recognize Text API

- Compare Speed
- Compare Accuracy

Chosen OCR Algorithm: Recognize Text API

Compile Final Script



Results & Discussions

Custom Vision Model

Table 1.0: Performance of Custom Vision Model

No. of Labelled Images	386
No. of Negative Images	572
Training Time	~10 mins
Precision	100.0%
Recall	93.5%
mAP	96.1%

Example of Successful Attempt

Figure 1.0: Raw Image

Figure 1.1: Crop of Predicted Bounding Box

Figure 1.2: Further Processed Image Sent to Recognize Text API

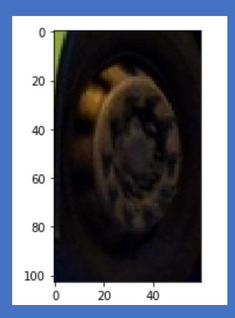
Response from Recognize Text API:

12

Example of Unsuccessful Attempt

Figure 2.0: Raw Image

Figure 2.1: Cropped Image of Predicted Bounding Box



Conclusion

We have implemented a proof of concept for the use of object detection, OCR and speech synthesis algorithms in the use bus number identification to aid the visually impaired in using public buses.

Future Work:
1. Incorporation of software into standalone wearable device
2. Implementation for different transport types
3. Extensions to VI locate key features of bus stop (bench, bollard)

References
[1] Custom Vision | Microsoft Azure (2019, December 09). Retrieved from https://azure.microsoft.com/en-us/services/cognitive-services/custom-vision-service/
[2] Computer Vision | Microsoft Azure (2019, December 09). Retrieved from https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/
[3] Text to Speech | Microsoft Azure (2019, December 09). Retrieved from https://azure.microsoft.com/en-us/services/cognitive-services/text-to-speech/

Figures:
Figure 1.0 taken myself, Figure 1.1, 1.2 derived from Figure 1.0.
Figure 2.0 taken from "https://businterchange.net/busphoto/albums/userpics/10001/SG5705E_883A.JPG". Figure 2.1 derived from Figure 2.0
Icons taken from "https://www.flaticon.com/free-icon/azure_873107", "https://miro.medium.com/max/725/1*FogMIj4gYwp3fTHLZuwavQ.png" and the Microsoft Azure Cloud and AI Symbol / Icon Set (https://www.microsoft.com/en-us/download/details.aspx?id=41937).