Research Plan

Rationale

Singapore aims for a public transport system that is inclusive to all, however, the visually impaired (VI) often face difficulties in using the public transport system. Advances in computer vision technology could potentially address this issue.

This project aims to create an all-in-one system to help the VI determine what bus is arriving at the bus stop using object detection, optical character recognition and text to speech techniques.

Research Question(s)

* What types of networks to use?
  + What are the limitations of each one?
  + How feasible are they to train and run (detection at real time speed)?
* How does one train and use these networks?
  + What can be automated via existing technologies?
  + What needs to be done from scratch?
* How effective will this solution be?
  + Will it be able to capture the buses each time?
  + Will it be user friendly enough to be used by the VI?

Hypothesis

We hypothesize that using modern machine learning techniques, we will be able to create a cohesive system that is able to detect the bus number of oncoming buses and relay this information to the visually impaired.

Engineering goal(s)

We aim to create a system that will take in a stream of images and determine the presence of buses in them and read off the bus number if a bus is present and then relay this information to the user in the form of audio.

Expected Outcome(s)

We hope that the developed system will be user friendly and accurate that the VI are able to reliably use it in their daily commutes. We also hope that it will prove the effectiveness of this approach towards enabling the VI to commute more independently and ganer the interest of bigger firms such as Microsoft and Google.

Procedures

We collect data by taking the perspective of the VI and filming the buses as they come and go at different bus stops at different times of the day. These videos are extracted into individual frames. We train an object detection network with this data and get it to identify the bus number location. We then use an OCR algorithm or neural network to extract the bus number from the location of the bus number specified by the object detection algorithm.

Risk and Safety

* Do not stand too close to the road while collecting data
* Ensure to not over spend credits on Azure or AWS if we use those cloud platforms

Methods for Data Analysis

Data Preparation

* Formatting and labelling the raw data
* Splitting the data for training, testing and validation

Training Phase

* Analysis of the performance of the neural networks
  + Accuracy
  + Recall
  + Precision

Bibliography from your literature review

"Azure for Students – Free Account Credit | Microsoft Azure." Microsoft, https://azure.microsoft.com/en-us/free/students/. Accessed January 06, 2020.

“Seeing AI | Talking camera app for those with a visual impairment." Microsoft, https://www.microsoft.com/en-us/ai/seeing-ai. Accessed January 06, 2020.

Vincent, James. "Google releases Lookout app that identifies objects for the visually impaired." The Verge, March 13, 2019, https://www.theverge.com/2019/3/13/18263426/google-lookout-ai-visually-impaired-blind-app-assistance.