Boston University
Electrical & Computer Engineering
EC463 Senior Design Project

First Prototype Testing Report

WhereTo

by
Team 5

Team Members:
John Burke - jwburke@bu.edu
Muhammad Ahmad Ghani - mghani@bu.edu
Haochen Sun - tomsunhc@bu.edu
Erick Tomona - erickshi@bu.edu
Marta Velilla - martava@bu.edu

Equipment:

Hardware:

- A personal computer

Software:

- Python3 Flask API
    - Image gathering
        - Google Street View Static API
    - Image text processing
        - Google Cloud Vision API
- Web Browser
    - Accessing locally hosted API via URL

Testing Procedure:

Our test was done entirely with software. The only hardware that was required was a personal computer on which we ran our code. We launched our Python API script locally in a virtual environment in order to access the endpoint using a web browser on the same machine. This allowed us to run the Python API without having to host it on the cloud or some other means.

Our API was designed to take as inputs two coordinates: an initial and a final coordinate. It progresses between these two coordinates and takes a set of pictures by querying Google's Street View Static API every few meters. The API is designed to take 8 images at each location, at intervals of 45 degrees. After collecting the images between the two supplied coordinates, the images are fed into Google's Cloud Vision API in order to have any text present within the images be read and output locally to a JSON formatted file.

Measurements:

The following is what we considered to be our measurable criteria for this prototype test:

I. Whether the Python API proves accessible locally, verified via the health endpoint
II. Whether the Python API captures and stores into a specified directory sets of images between the specified/supplied coordinates. These images should be in sets of 8, as they are taken every 45 degrees for an entire rotation.
III. Whether the Python API outputs a JSON formatted file containing the results of processing the images gathered between the two coordinates through Google Cloud Vision Text analysis.
    A. For every image processed between the provided coordinates, the JSON file will have a distinct section detailing the text extracted from that specific image.

Here are the results from our testing in regards to these criteria:

I. The Python API did prove accessible using the health endpoint
II. The Python API successfully captured the images expected between the two supplied coordinates. It was also successful in that for each location, a set of 8 pictures taken at each different angle was provided by querying the Google Street View Static API.
III. The Python API successfully outputs the expected JSON file. This file contained the text information that the Google Cloud Vision text analysis tool API was able to see. While this step was successful, the text was read from all parts of the image, not just parts relevant to parking signage.

Conclusion:

Our tests from this prototype show that we are on the right track with our implementation and have a clear way forward to improve our design. We know that we are able to construct a working Python API as utilization of our health endpoint was successful during our test in the lab. Additionally, we have shown competency in our utilization of Google API services as the end-to-end process is able to run without error and store and provide annotations for the images in a JSON file.

The path forward for our design will be to implement Google Cloud Vision or another image text analysis model in conjunction with a model to detect and bind parking signage. A significant enhancement in this direction is the integration of the YOLO (You Only Look Once) model for object detection. We have sourced a pre-trained dataset online to train the YOLO algorithm specifically for identifying parking signs. By leveraging YOLO's robust and efficient object detection capabilities, we aim to improve the precision of bounding parking signs in images This will allow us to read text only that relates directly to parking, or other important signage. We believe this is a good step moving forward as it will make it so that we receive less irrelevant text information from our calls to the Google Cloud Vision model. Signage regarding anything but parking rules can be ignored by our model, so implementing this next step of road sign detection is crucial moving forward.