HailRick: A camera tool to help the visually impaired in India hail auto-rickshaws

Sriraman Madhavan
Department of Statistics, Stanford University
sriraman@stanford.edu

Abstract

The visually impaired in India face several challenges in using public transport. Hailing auto-rickshaws is a major obstacle for the visually impaired, because the steps typically involve sighted assistance, and a study identified perpetual dependence on sighted assistance for external travel a leading cause of frustration for a large number of visually impaired people. This project aims to use Computer Vision to eliminate the need for sighted assistance in hailing auto-rickshaws, by building a tool that uses the smartphone camera to alert the visually impaired user when an auto-rickshaw is passing by. The image classification model to detect an auto-rickshaw in a given image achieved an accuracy of 95% on a balanced test dataset.

1 Introduction

Auto-rickshaws are motorized three-wheeler taxis, which are ubiquitous in Indian cities for public transport. A typical Indian city has around 60,000 auto-rickshaws running every day. Hailing an auto-rickshaw typically involves the following steps: standing on the curb of a busy street, spotting an empty auto-rickshaw passing by, sticking the arm out, and shouting "Auto!" to alert the driver. Hence, hailing auto-rickshaws is a major obstacle[1] for the visually impaired, as it typically needs sighted assistance. A study identified perpetual dependence on sighted assistance for external travel a leading cause of frustration for a large number of visually impaired people.[2] This project aims to use Computer Vision to eliminate the need for sighted assistance in hailing auto-rickshaws.

Specifically, this project aims to solve the problem faced by the visually impaired in spotting an auto-rickshaw, by building a tool that uses the smartphone camera to alert the visually impaired user when an auto-rickshaw is approaching. Smartphone adoption among visually impaired people in India is gaining traction[3] [4], and offers promising solutions to several problems faced by the visually impaired. This tool would use deep learning to detect an auto-rickshaw on the street, and will eliminate a major obstacle contributing to the need of sighted assistance in using auto-rickshaws. A typical use of the tool would involve the user standing on the curb of a busy street, pointing the smartphone camera at the road, and the phone would vibrate when an auto-rickshaw is approaching or passing by.

---

[1] https://trid.trb.org/view/890862
[2] https://escholarship.org/uc/item/5pv2k256
[3] http://www.nixdell.com/papers/vism_assets_camera_ready.pdf
[4] http://joyojeet.people.si.umich.edu/wp-content/uploads/2017/03/Smartphone-Bangalore-Seoul.pdf

2 Related works

Previous attempts to make public transport more accessible to the visually impaired in India have typically involved intrusive solutions (such as installing devices on vehicles), which are not scalable. Computer vision methods have been proposed for bus route number recognition in Taiwan[5] and Thailand[6]. There has been no work done to make auto-rickshaws more accessible to the visually impaired, even though a typical Tier 1 city in India has more than 50,000 auto-rickshaws running every day.[7]

3 Dataset

The specific task is to classify a given image (periodic snapshots taken via camera) as having an auto-rickshaw or not. There are no readily available datasets having images of auto-rickshaws. Also, the evaluation should ideally happen on snapshots taken from videos of Indian roads. Hence:

- Training:
  - 990 images of auto-rickshaws downloaded via Google Images.
  - 980 images of empty roads, cars and bikes on the road, downloaded via Google Images.

- Dev:
  - 150 screenshots of auto-rickshaws on the road, taken from roughly 20 different YouTube videos of busy Indian roads.
  - 150 screenshots of empty roads, cars, bikes and buses taken from the same videos.

- Test:
  - Similar to Dev.

Typical training set images of an auto-rickshaw look like this:

Typical dev/test set images of an auto-rickshaw look like this:

[5] https://link.springer.com/chapter/10.1007/978-3-319-05458-2_32
[6] http://ieeexplore.ieee.org/document/7748910/?reload=true
[7] http://wricitieshub.org/charts-graphs/market-size-auto-rickshaws-indian-cities

4 Methods

The specific task is to classify a given image as having an auto-rickshaw or not. As a baseline, a logistic regression model was implemented, followed by training a shallow neural network, a deep neural network and finally a convolutional neural network to compare and understand the results.

5 Results

Model	Train accuracy	Dev accuracy	Test accuracy
Logistic Regression	76%	69%	67%
Feed-forward Neural Network	93%	89%	87%
Convolutional Neural Network	98%	96%	95%

The baseline model – logistic regression, could achieve a test accuracy of 67%, while neural networks performed much better. The final convolutional neural network achieved a test accuracy of 95%. Error analysis indicates that the model has problems in identifying auto-rickshaws when the lighting is poor, when the rickshaw is not of the typical auto-rickshaw color (black + yellow/green) and when other vehicles are of the typical auto-rickshaw color.
Code: https://github.com/sriram172/hailrick

6 Future Work

Future work involves and is not limited by:

- Improving the test accuracy: Currently, 1 in 20 predictions would be wrong, and that's unreasonable for this application. Improving the CNN model further is necessary.
- Implementing SqueezeNet to shrink the model to fit on a typical Indian smartphone.
- Building a mobile app that integrates this model with the phone camera.