Intelligent Mobility Aid for Visually Impaired

Shantanu Deshmukh

Akhilesh Jichkar

Saketh Saxena

The problem – Visual Impairment

- The affected - 253 million people visually impaired, 36 million fully blind [1].
  - Many from low income backgrounds
  - Require mobility aids to navigate
- Existing mobility aids
  - Eg. Canes, guide dogs, “mini-guides”
  - Expensive, bulky, restrictive and require training.
- Need for an inexpensive and constantly evolving(intelligent) solution

Our approach

- An android application to aid in navigation by detecting obstacles combining
  - Phone's Portability
  - Camera features for input
  - Image classification using CNNs – TensorFlow lite
  - **Obstacles** – bike, newspaper vending machine, people, bench, trashcan
- Advantages over existing system
  - Inexpensive + easily accessible
  - Lightweight - phone's keep getting thinner
  - Extensible – easily train new obstacles

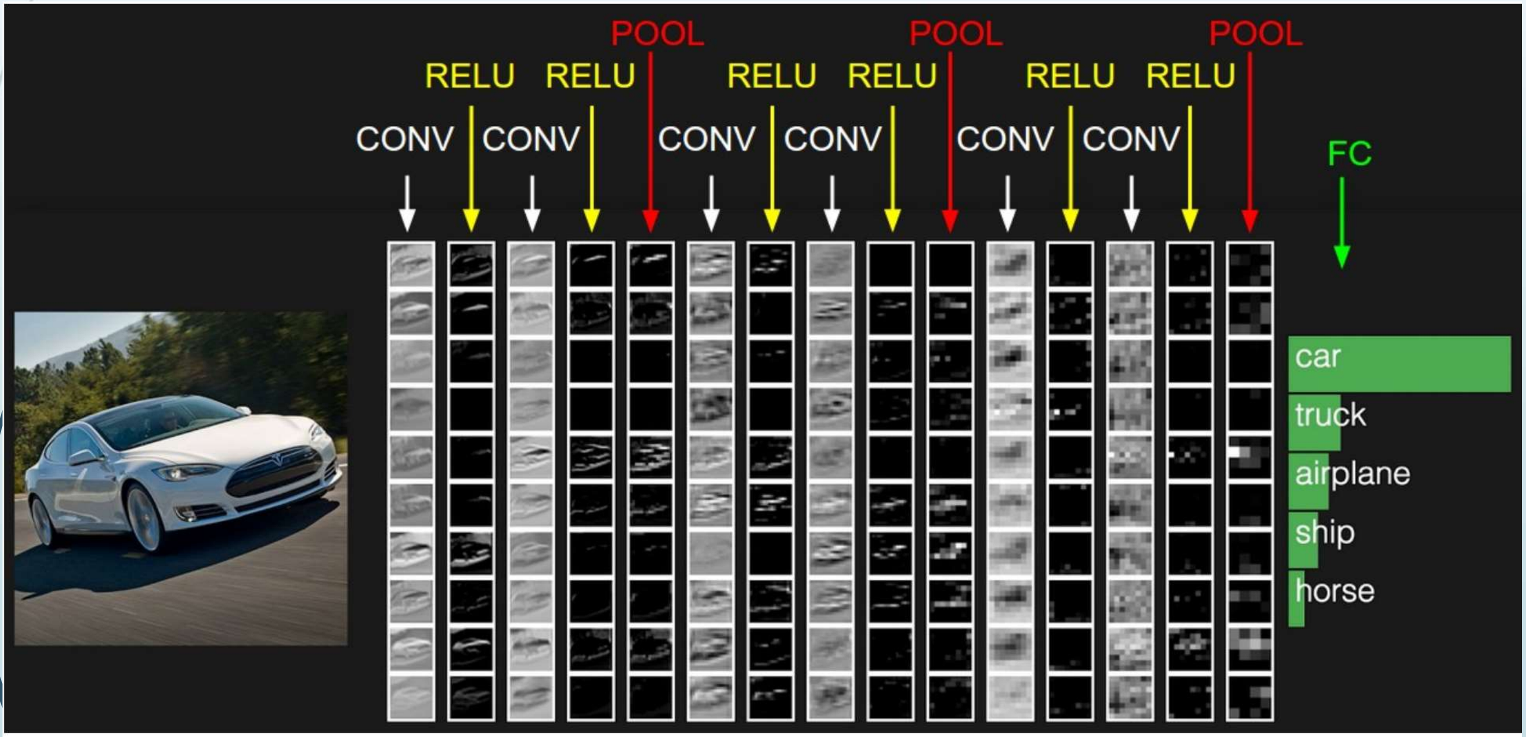
Image Recognition using CNN

Convolution Layer
convolution filters
single valued output feature map

ReLU (Rectified linear unit) Layer
Thresholding at zero

Pool Layer
Reduces Dimensionality

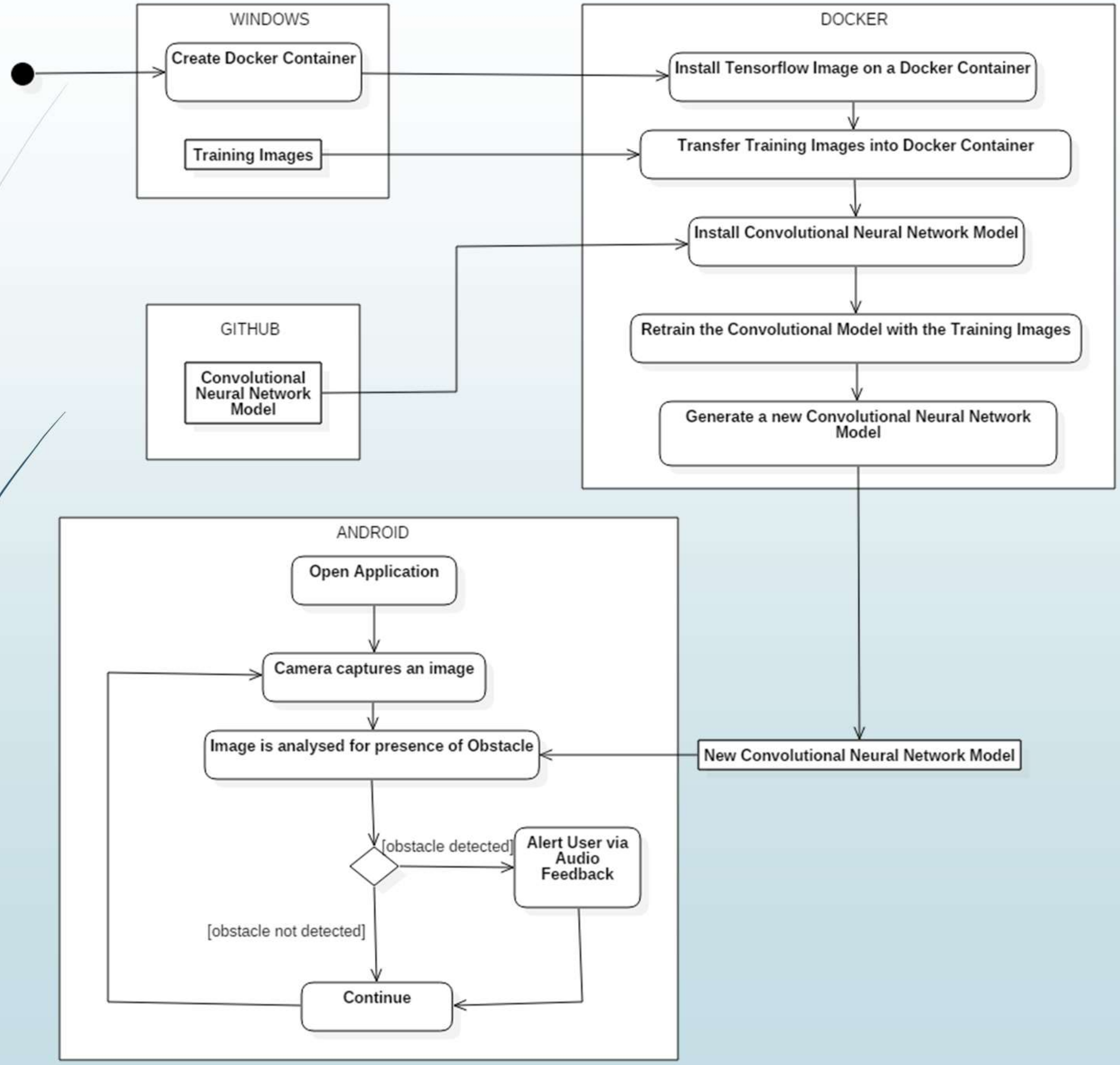
Dense Layer
Final classification
Fully Connected



TensorFlow Lite (mobileNet)

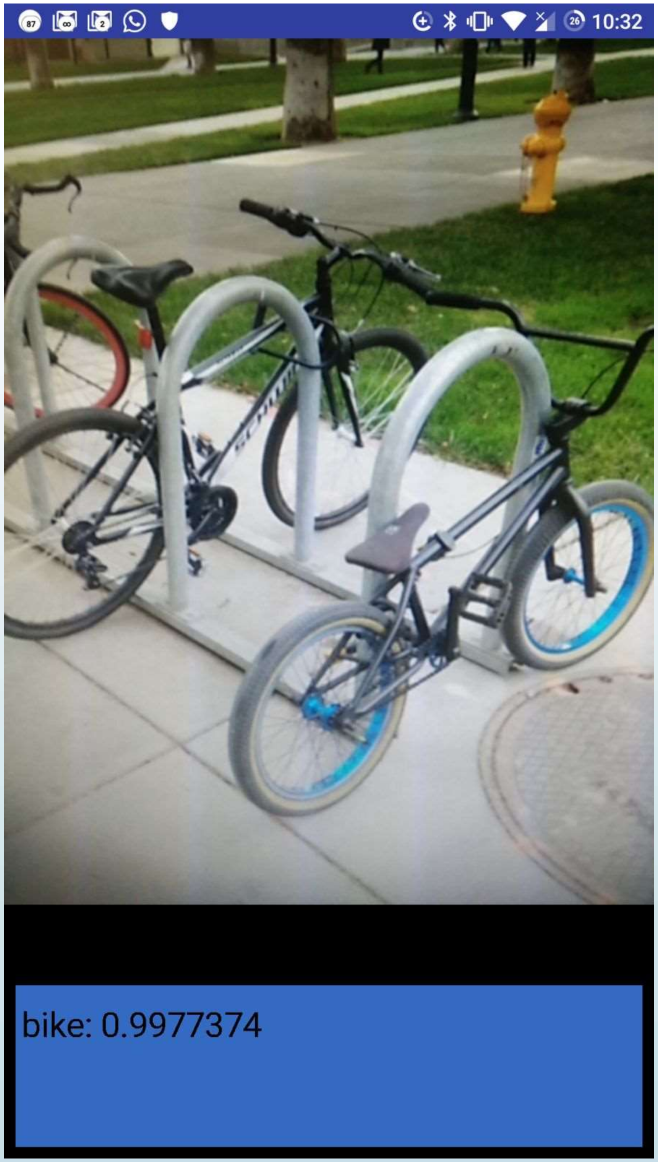
- Open source software library for numerical computation on multidimensional arrays.
- Lightweight solution for mobile
- Ability to serve ‘offline’ use cases
- MobileNet(Pre-tested model):
  - Mobile-first computer vision model
  - Effective on small, low-latency, low-power devices
  - Smaller but lower in accuracy

System Architecture



User Interface

Showing the app detecting a bike in the immediate vicinity.



Performance Metrics

- Agent
  - Confidence Score

    Summation of values obtained at the last layer of the CNN for each label/class in Fully Connected Dense Layer
  - Prediction Accuracy

    (Correct Predictions / Total Sample Data) * 100

- Android Application
  - Average latency

    Average time between detecting the obstacle and returning output

Experimental Results – Prediction Accuracy

Tested against 25 images per obstacle

Obstacle	Correct Predictions	Prediction Accuracy
Bench	18	72.0 %
Newspaper vending machine	25	100.0 %
Person	15	60.0 %
Trashcan	21	84.0 %
Bike	21	84.0%

Experimental Results- Confusion Matrix

Predicted ➡ Real ⬇	Obstacle Detected	Obstacle not Detected
Obstacle Present	100	25
Obstacle not Present	25	400

- Accuracy: 90.9 %
- Precision: 80.0 %
- Recall: 80.0 %
- Specificity: 94.1 %
- NPV: 94.1 %

Experimental Results – Average Latency

Average latency between image capture, inference and output is 203.4 ms for all obstacles.
Thus, as per our criteria our platform is strong.

Obstacle	**Average Latency *(milliseconds)***
Bench	215
Newspaper vending machine	185
Person	212
Trashcan	198
Bike	207

Phone: Xiaomi Redmi Note 4

Demo Video

Challenges & Future Work

- Challenges
  - Camera’s field of view is limited and cannot perceive the depth of an image.
  - Training the model on all possible obstacles is not practically feasible.
- Future Work
  - Multiple obstacle detection needs to be implemented if more than one obstacle is present in the frame.
  - Distance of obstacles from the user
  - Identify known people
  - Obstacles in context of the Image

References

[1] WHO, “Vision Impairment and blindness factsheet”
http://www.who.int/mediacentre/factsheets/fs282/en/