

# Public Transport Access Detection

Dragoş-Mihail Comănac

# Contents

<b>Abstract</b>	<b>4</b>
<b>1 Introduction</b>	<b>5</b>
1.1 Context and motivation . . . . .	5
1.2 Objectives . . . . .	5
1.3 Paper structure . . . . .	5
<b>2 State of the art</b>	<b>6</b>
<b>3 Theoretical foundations</b>	<b>7</b>
<b>4 Design and implementation</b>	<b>8</b>
<b>5 Experimental results</b>	<b>9</b>
5.1 Dataset . . . . .	9
5.2 Hyperparameter tuning . . . . .	9
5.3 Performance evaluation . . . . .	9
5.4 Comparison with other methods . . . . .	9
<b>6 Conclusions and future work</b>	<b>10</b>
<b>7 Bibliography</b>	<b>11</b>

# Application and original contribution

## Classification

**AMS Mathematical Subject Classification - 68T45 Machine vision and scene understanding**

**ACM Computing Reviews Categories and Subject Descriptors - I.4 IMAGE PROCESSING AND COMPUTER VISION - I.4.6. Segmentation - Edge and feature detection**

## Application lifecycle

1. Application is opened
2. User chooses auditive or visualization mode using a voice command
3. In auditive mode, the live object detection automatically starts, and information is provided to the user in an auditive manner
4. In visualization mode the user can load an image and perform detection on it or can start object detection on the live feed from the camera
5. Application is closed by voice command or through a button

## Functionalities

- choose auditive mode or visualization mode using a voice command
- auditive mode: perform live object detection on cars, busses, and license plates and based on the results of detection, provide auditory clues that guide the user towards the detected car or bus. If the detected object is a bus, then provide information about the location of the doors and the line of the bus (if available). If the detected object is a car, then provide information about the location of the doors and the license plate (if available)

- visualization mode: display the results of object detection on an image or on the live feed from the camera
- close the app using a voice command

## Original contribution

The application aims to help the visually impaired persons by using computer vision and line detection, on a mobile device, to provide spatial information about public transport such as busses or ride sharing access. An object detector model will extract information about the vehicle and a line detector will analyze the part of the image containing the vehicle to extract information about the doors of the vehicle. This should eliminate the need for a dataset that has specific bounding box annotations for the doors.

# Abstract

- short description of the approach and results

# 1. Introduction

## 1.1 Context and motivation

## 1.2 Objectives

## 1.3 Paper structure

## 2. State of the art

- discussion on the current methods used in object detection and line detection

### 3. Theoretical foundations

- description of the basic concepts that are used



## 4. Design and implementation

- description of the proposed solution

## 5. Experimental results

### 5.1 Dataset

### 5.2 Hyperparameter tuning

### 5.3 Performance evaluation

### 5.4 Comparison with other methods

## 6. Conclusions and future work

- summary of the solution - critical analysis of the solution - future improvements

## 7. Bibliography

- [1] Cuneys Akinlar and Cihan Topal. Edlines: A real-time line segment detector with a false detection control. *Pattern Recognition Letters*, 32:1633–1642, 10 2011.
- [2] Md. Zahangir Alom, Tarek M. Taha, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C. Van Essen, Abdul A. S. Awwal, and Vijayan K. Asari. The history began from AlexNet: A comprehensive survey on deep learning approaches. *CoRR*, abs/1803.01164, 2018.
- [3] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *CoRR*, abs/2004.10934, 2020.
- [4] Lamya El Alamy, Sara Lhaddad, Soukaina Maalal, Yasmine Taybi, and Yassine Salih-Alj. Bus identification system for visually impaired person. In *2012 Sixth International Conference on Next Generation Mobile Applications, Services and Technologies*, pages 13–17, 2012.
- [5] Ross B. Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *CoRR*, abs/1311.2524, 2013.
- [6] Lilit Hakobyan, Jo Lumsden, Dymrna O’Sullivan, and Hannah Bartlett. Mobile assistive technologies for the visually impaired. *Survey of Ophthalmology*, 58(6):513–528, 2013.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *CoRR*, abs/1406.4729, 2014.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [9] Derek Hoiem, Yodsawalai Chodpathumwan, and Qieyun Dai. Diagnosing error in object detectors. *Computer Vision ECCV 2012*, pages 340–353, 2012.

- [10] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, and Kevin Murphy. Speed/accuracy trade-offs for modern convolutional object detectors. *CoRR*, abs/1611.10012, 2016.
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.
- [12] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Mallocci, Alexander Kolesnikov, and et al. The open images dataset v4. *International Journal of Computer Vision*, 128(7):1956–1981, Mar 2020.
- [13] Dongjin Lee, Hosub Yoon, Chankyu Park, Jaehong Kim, and Cheong Hee Park. Automatic number recognition for bus route information aid for the visually-impaired. In *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 280–284, 2013.
- [14] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. Ssd: Single shot multibox detector. *Lecture Notes in Computer Science*, page 21–37, 2016.
- [15] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640, 2015.
- [16] Joseph Redmon and Ali Farhadi. YOLO9000: better, faster, stronger. *CoRR*, abs/1612.08242, 2016.
- [17] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *CoRR*, abs/1804.02767, 2018.
- [18] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015.
- [19] Na Wang, Ruoyan Chen, and Kang Xu. A real-time object detection solution and its application in transportation. In *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, pages 486–491, 2021.
- [20] Junjie Yan, Zhen Lei, Longyin Wen, and Stan Z. Li. The fastest deformable part model for object detection. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2497–2504, 2014.