



# Technologies for Safe & Efficient Transportation

THE NATIONAL USDOT UNIVERSITY  
TRANSPORTATION CENTER FOR SAFETY

Carnegie Mellon University

UNIVERSITY of PENNSYLVANIA

---

## Automatic Counting of Pedestrians and Cyclists

---

### FINAL RESEARCH REPORT

Bernardo R. Pires (PI), Jian Gong, Chris Kaffine,  
Mehmet Kemal Kocamaz, John Kozar, Ganesh Kumar  
Nunnagoppula, Dhruv Saksena

Contract No. DTRT12GUTG11

## **DISCLAIMER**

**The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.**

## Introduction and Problem Description

On September 2014, U.S. Secretary of Transportation Anthony Foxx unveiled a new Transportation Action Plan [1] to increase active, no-motorized transportation. On it, the Department of Transportation reinforced its commitment to “making safe walking and biking a reality.” Although the health and environmental benefits of a non-automobile commute are well known, it is still difficult to understand how to get more people to take up active transportation. Infrastructure can have a dramatic effect on cycling and walking adoption, but represents a significant outlay of government resources. Thus, concrete usage statistics are paramount for assessing and optimizing such spending. The goal of this project is to provide actionable data for government officials and advocates that promote bicycling and walking. The project created a vision-based cyclist and pedestrian counting system that allow for automatic and human-assisted data collection and analysis. Unlike traditional non-vision counting methods, our system has the potential for much higher accuracy while providing valuable usage and demographic data that simply cannot be collected by other sensors.

## Approach

The pedestrian and cyclist counting project was a result of a real-world need from the City of Pittsburgh to determine the usage of newly-created dedicated bike lanes throughout the city. Due to the relatively large area of bike paths for which it would be desirable to obtain information, a portable data collection system was deemed the most effective solution. The developed data collection device (Figure 1) consists of a ruggedized Windows tablet (Panasonic Toughpad), an extensible pole, and a miniature bullet camera. In order to collect data, the bullet camera is mounted at the top of the pole, which is extended to a suitable height. The whole system is fastened to a lamp post or other sturdy city fixture. The tablet is used to verify that the camera is pointed accurately at the bike lane and to control the data collection. The system is battery powered and allows for the collection of up to 12 hours of data on a full charge.



**FIGURE 1.** Portable data collection system

### Collected Data

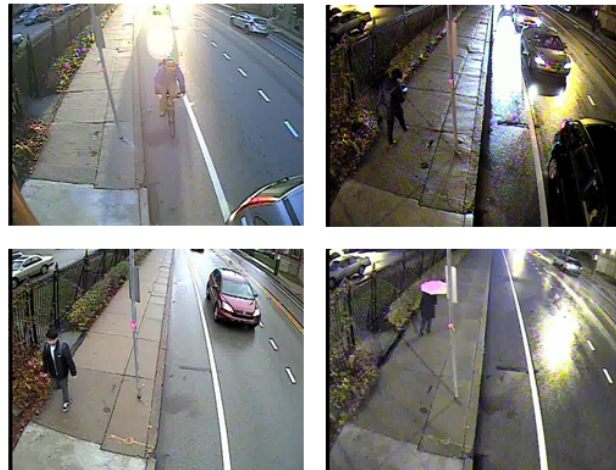
Two datasets were collected during the project: The Schenley Dataset was used to develop the counting system. It was collected during a period of 11 days during the month of June 2015, usually from 9am to 4pm. Data was collected at two locations alongside a bike lane on Schenley Drive, next to CMU's campus in Pittsburgh. Data was recorded at different view-points, weather, and illumination conditions. Approximately 50 hours of video were recorded. A sample of the dataset is shown in Figure 2. Ten hours of the recorded dataset were labeled. The labeled dataset includes 541 unique pedestrians and 111 unique cyclists. In total, ground truths were generated for approximately 270,000 separate pedestrian observations and approximately 49,000 separate bicyclist observations.



**FIGURE 2.** Samples of Schenley Dataset

Data was recorded at different view-points, weather, and illumination conditions. Approximately 50 hours of video were recorded. A sample of the dataset is shown in Figure 2. Ten hours of the recorded dataset were labeled. The labeled dataset includes 541 unique pedestrians and 111 unique cyclists. In total, ground truths were generated for approximately 270,000 separate pedestrian observations and approximately 49,000 separate bicyclist observations. All the videos for this dataset focus on the same bike lane but in different directions. Among the 11 days recorded, 5 days of data face the bike lane in one direction and all the other days facing in the opposite direction. Since the system is disassembled at the end of each day, differences are present in the camera height and angle from day to day. The weather varies from clear, to cloudy to rainy, often during the same day which leads to large background brightness variations throughout the dataset.

The Bayard dataset was collected to provide actionable information to the City of Pittsburgh. It was collected at the intersection of Craig Street and Bayard Street in Pittsburgh during a period of 7 days in December of 2015. The dataset aimed to collect specifically the rush hour, from 7am to 9am in the morning and from 4pm to 6pm in the afternoon. Approximately 24 hours of video were recorded. Due to technical difficulties, the data was not recorded in the afternoon of 12/2 and in the morning of 12/4. Sample images from the recorded data can be seen in Figure 3. All recordings were captured approximately from the same direction, facing west on Bayard Street. The recorded dataset includes challenging illumination and weather conditions which make the counting process difficult for the automated systems. Figure 3 displays some of these challenges, such as the head lights of the cars, rain, insufficient lighting, shadows of the objects, and different view-points.



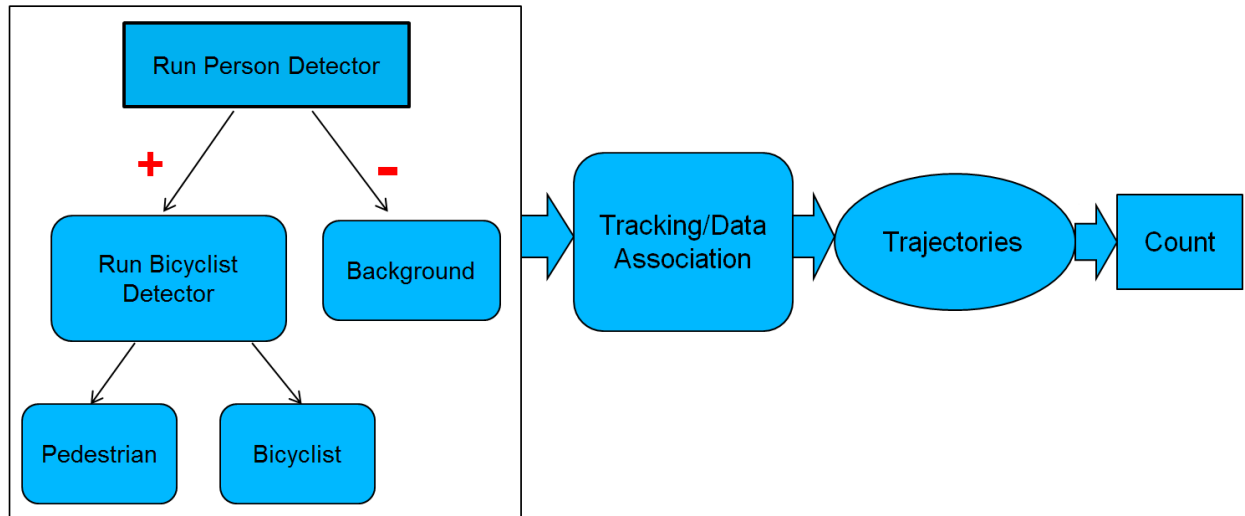
**FIGURE 3.** Samples of the Bayard Dataset

## Methodology

Object (vehicles, people, bicyclists, animals, etc.) counting is one of the essential tasks in Computer Vision systems for multiple applications and domains. This is a challenging and open problem due to several factors, including illumination variations, pose changes, appearance similarities, and occlusion between multiple objects and between the objects and the environment. There are two primary approaches to this problem: (1) Region of Interest (ROI) counting [2,3,4]; and (2) Line of Interest (LOI) counting [5,6,7]. The goal of ROI counting is to estimate the number of the objects in a desired region of the image at a specific time. In the LOI counting approach, an object is counted if it crosses a virtual line of interest in the image. This project uses the LOI approach, as it better matches the statistics desired by the end users of the system (City of Pittsburgh).

The developed pedestrian and bicyclist counting method follows a detect-track-count approach. The detection step of the algorithm consists of a cascaded classifier. In the first level of the cascaded classifier, the pedestrians and the bicyclists are detected. In the second level of the classifier distinguishes the bicyclists from the pedestrians. The cascaded classifier is based on the state-of-art object detector in [8]. In order to reduce the false alarms of the original detector, it

has been extended by incorporating additional cues. Namely, foreground segmentation, geometry prior information, and temporal moving direction are incorporated into the non-maximal suppression step of the algorithm. The final detections are obtained from by using a multi-cue clustering algorithm. The overall detector, including the priors and the multi-cue clustering, is referred to as the Extended Detector. A data association method employs the output of this detector to produce trajectories which are ultimately used to determine the pedestrian and cyclist counts. The overview of the proposed method can be seen in Figure 4.



**FIGURE 4.** Overview the proposed counting approach. In the first step, a cascaded classifier detects pedestrians and bicyclists at each video frame. The subjects are tracked by associating the detections with the trajectories. In the last step, the subjects are counted if their trajectories cross a virtual line in the image.

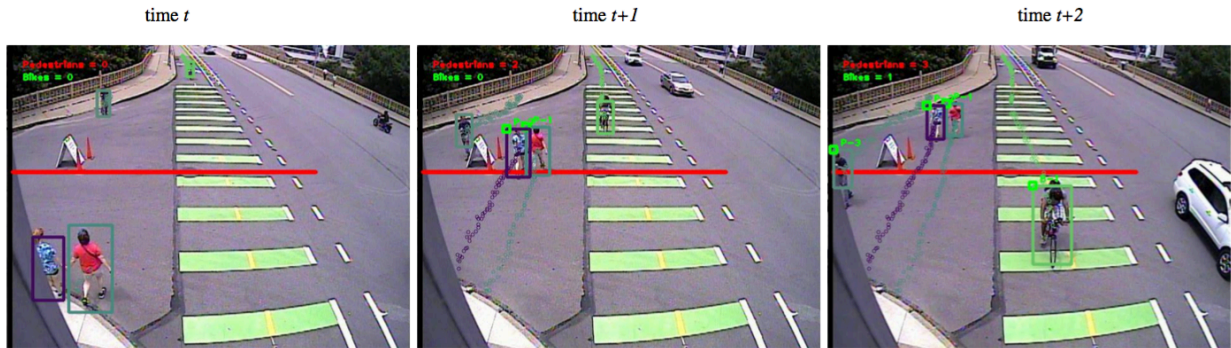
The Schenley Dataset was used to train and evaluate the developed method. In order to train the proposed cascaded detector, the dataset was divided into a train dataset (Schenley-Train), and a test dataset (Schenley-Test). The Schenley-Train dataset consists of 40k pedestrians, 13k bicyclists, and 200k negative samples. The Schenley-Test dataset was formed in a way that no same pedestrian or bicyclist exist in Schenley-Train. Schenley-Test includes 10k pedestrian, 5k bicyclists, and 10k negative samples. The training of the first and second levels of the cascaded detector took approximately 18 hours and 6 hours, respectively, on a machine which has an Intel i7-3930K 6 core CPU processor and a 1536 core NVIDIA GPU.

## Real World Test and Evaluation

Figure 5 shows sample results for the Schenley Dataset. The confusion matrix for this dataset can be seen in Table 1. The Extended Detector achieved 95.1%



accuracy in the classification of the samples. Total of 461 background samples were labeled as the person and 421 pedestrian samples were evaluated as the background by the detector. Only 11 pedestrians were classified as bikes in the second level of the detector which shows that it has high accuracy to distinguish between the pedestrians and bicyclists. Some false detections of the detector can be seen in Figure 6. They usually occur when the objects are far from the camera, and/or are partially occluded.



**FIGURE 5.** Sample counting results for the Schenley Dataset. Bikes and pedestrians are counted when they cross the red line.

	Peds	Bikes	Background
Predicted Peds	9568	4	453
Predicted Bikes	11	4665	8
Predicted Background	421	331	9539

**TABLE 1.** Results of the Extended Detector on the Schenley Dataset



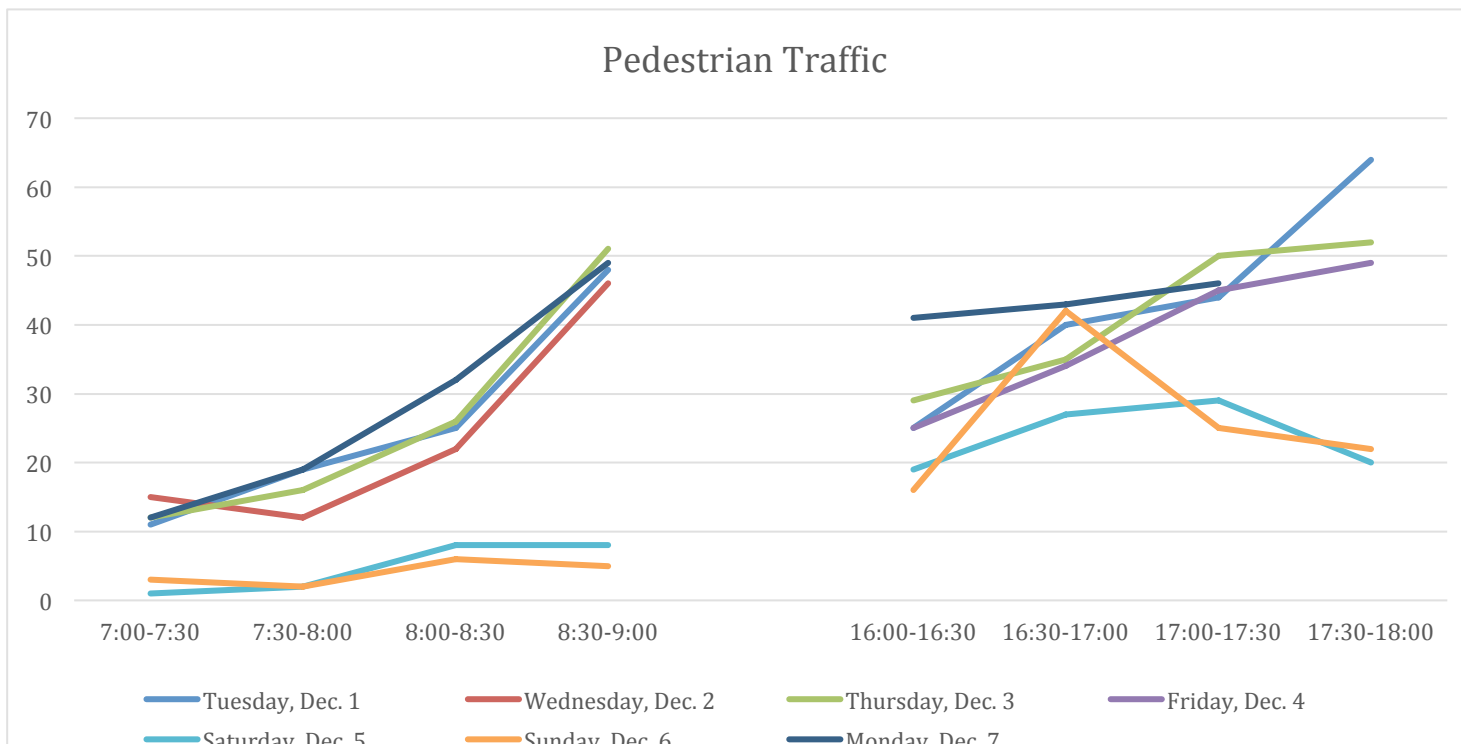
**FIGURE 6.** Examples of false detections of the cascaded detector. Column headings show the classification results of the detector. Low resolution of far objects, and/or partial occlusion are main reasons of the failure cases.

The method developed on the Schenley Dataset was applied to the Bayard Dataset. Because the counting results for Bayard were to be shared with the City of

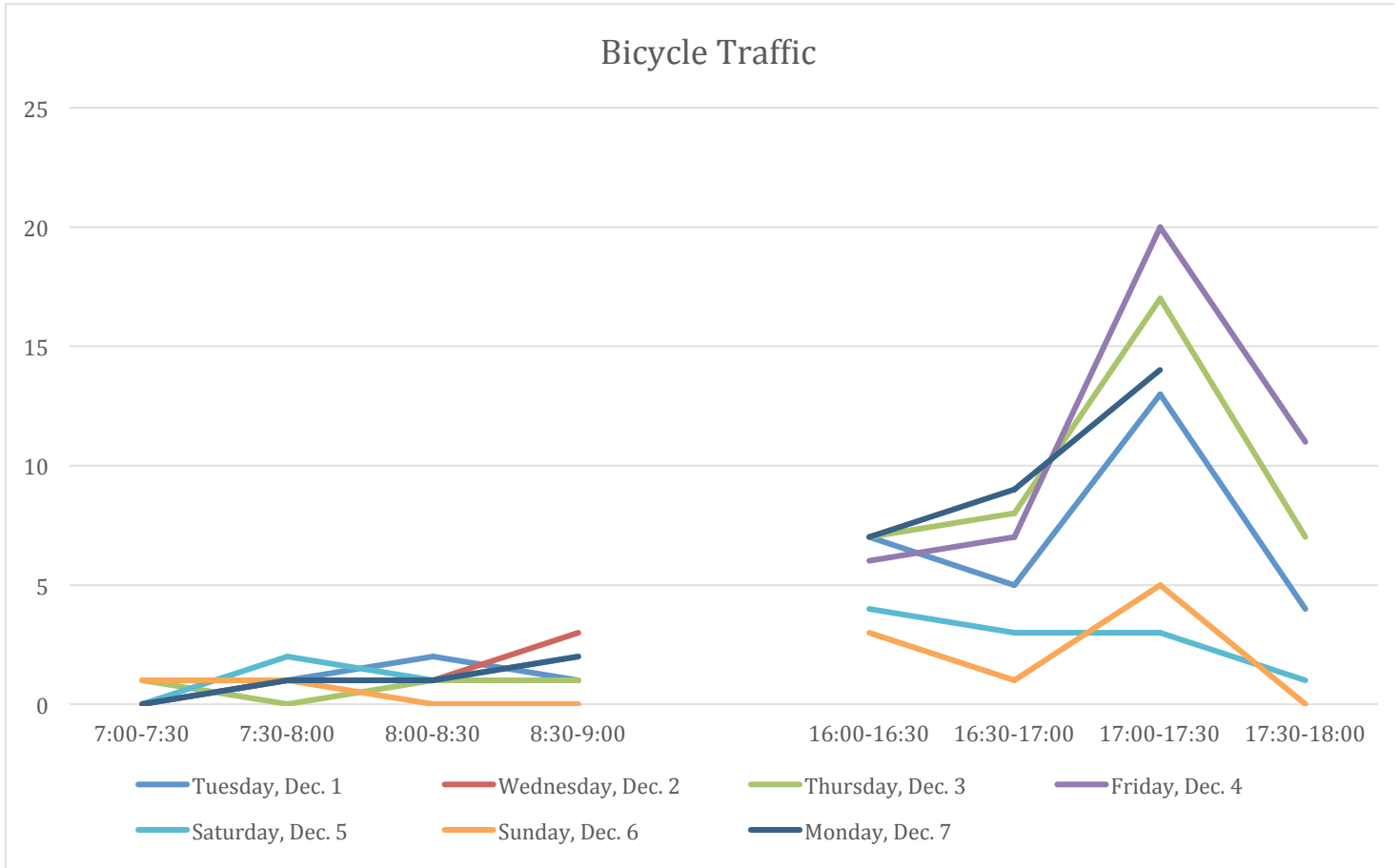
Pittsburgh, possibly to make policy decisions, all counts were verified manually. The number of pedestrians and cyclists for each day can be seen in Figure 7. In order to observe the hourly usage, the subjects were counted for each half hour of the recorded time. Some anomalies pedestrians and cyclists were observed in some days, such as a pedestrian walking in the bike lane, or cyclists on the walkway. Table 2 lists these anomalies, which can be used to investigate the need for corrective action by city officials, such as additions of road signs, changes to road layout, or even the creation of a protected lane.

Date	Time	Anomaly
12/5/2015	7:30-8:00	Pedestrian in bike lane
12/5/2015	8:30-9:00	Bike moving in rightmost lane; Pedestrian in bike lane; Bike on sidewalk
12/5/2015	16:00-16:30	Bike on sidewalk
12/5/2015	16:30-17:00	White car enters bike lane
12/5/2015	17:00-17:30	Two pedestrians run across road, Person crosses road from sidewalk
12/5/2015	17:30-18:00	Bike on road (twice), Person crosses road from sidewalk
12/6/2015	7:00-7:30	Pedestrian crosses road
12/6/2015	17:30-18:00	Person walks across road onto sidewalk
12/7/2015	7:00-7:30	Person walks on bike lane, Person crosses road
12/7/2015	7:30-8:00	Person gets on sidewalk from road
12/7/2015	8:00-8:30	Person on bike lane + crosses road
12/7/2015	8:30-9:00	Black Chevy in bike lane, Man with suitcase crosses road

**TABLE 2.** Anomalies observed on the Bayard Dataset







**FIGURE 7.** Pedestrian and Bicycle Traffic for the Bayard Dataset

Regarding pedestrian data, it appears that the morning rush is later than regular automotive traffic, as the number of pedestrians is still increasing at 9am. This might be related to the proximity to the universities and the large student population. It is also interesting that there is little morning pedestrian traffic on weekends. Afternoon pedestrian traffic seems to not vary much throughout the week.

Regarding bike traffic, and taking into account that we are only observing traffic in the eastbound direction, the data seems to support the notion that most cyclists using this bike lane take the east-west direction in the morning and the reverse direction in the afternoon. This, again, supports the notion that cyclists using this route commute into Oakland (and the University of Pittsburgh in particular) from the western neighborhoods. It is also interesting to note that the bike data has a clear spike at 5pm on weekdays. The diminished traffic on

weekends also supports the notion that the bike lane is used mostly for commuting and less for recreation.

Regarding the observed anomalies, most included jaywalking, although there are instances of bikes on sidewalk and on road instead of bike lane. Also, there are multiple cases of both pedestrians and cars entering the bike lane.

## Conclusions and Recommendations

The project developed a system for automatic and human assisted counting of pedestrians and bicyclists. The feedback from the City of Pittsburgh partners was very positive and plans are underway to collect and process even more extensive datasets during the 2016 calendar year. The method was also received very positively by the research community with a publication on the top-tier conference for practical applications of computer vision [9].

## References

- [1] **Safer People, Safer Streets: Summary of U.S. Department of Transportation Action Plan to Increase Walking and Biking and Reduce Pedestrian and Bicyclist Fatalities.** September 2014
- [2] Antoni B. Chan and Nuno Vasconcelos, **Counting People with Low-Level Features and Bayesian Regression**, IEEE Transactions on Image Processing, Volume 21, Number 4, Pages 2160-2177, 2012.
- [3] Jingwen Li, Lei Huang, and Changping Liu, **Robust people counting in video surveillance: Dataset and System**, Advanced Video and Signal-Based Surveillance, Pages 54-59, 2011.
- [4] D. Ryan, S. Denman, C. Fookes, and S. Sridharan, **Crowd Counting Using Multiple Local Features**, Digital Image Computing: Techniques and Applications, Pages 81-88, 2009.
- [5] Yang Cong, Haifeng Gong, Song-Chun Zhu, and Yandong Tang, **Flow mosaicking: Real-time pedestrian counting without scene-specific learning**, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Pages 1093-1100, 2009.
- [6] Zheng Ma and A. B. Chan, **Crossing the Line: Crowd Counting by Integer Programming with Local Features**, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Pages 2539-2546, 2013.
- [7] Kin-Yi Yam, Wan-Chi Siu, Ngai-Fong Law, and Chok-Ki Chan, **Effective bi-directional people flow counting for real time surveillance system**, IEEE International Conference on Consumer Electronics, Pages 863-864, 2011.

[8] Rodrigo Benenson, Mohamed Omran, Jan Hosang, and Bernt Schiele, **Ten years of pedestrian detection, what have we learned?**, Proceedings of the European Conference on Computer Vision (ECCV), Pages 613--627, 2014.

[9] Mehmet Kemal Kocamaz, Jian Gong, and Bernardo R. Pires, **Vision-based Counting of Pedestrians and Cyclists**, IEEE Winter Conference on Applications of Computer Vision WACV'16, March 7-9, 2016, Lake Placid, NY, USA.