2 Month PhD Plan

Morten Bornø Jensen

mboj@create.aau.dk

June 29, 2016

1. Project abstract

This project is about analyzing vulnerable road users(VRU) in the traffic scene using computer vision methods. The number of personal vehicles in Denmark has raised to 2.391.000 from 2015-2016 [23] which correlated with 9 out of 10 dane owning a bike [25] unfortunately presents a risk for accidents.

Traffic researchers have a desire to prevent these accidents and conflicts by capturing data at e.g. selected intersections. The occurrence of near-accidents or conflicts are a lot higher than actual accidents. A large amount of data is therefore needed in order to capture sufficient conflicts and accidents at a given site. This can easily add up to hundred of hours traffic data, which is a very time consuming process for the traffic researchers to manually look through.

Computer vision is an obvious tool to automate parts of this very time consuming process. It can be used to detect and classify various objects in the traffic scene, e.g. vehicles, cyclists, or pedestrians. In order to replace a large part of the task that is currently done manually by the traffic researchers, the tool must be able to this in all sorts of weather and lighting condition. Furthermore, the tool must be able to detect a variety of other parameters as interesting traffic studies, e.g. conflict studies, is usually a combination of events such as multiple objects crossing paths or red running.

In order to meet the criteria of doing data reduction in varied traffic data captured continuously throughout several weeks, different capturing modalities are used. The most common image acuition modality is RGB which is a good descriptor of a vehicle, cyclist, or pedestrian, but relies on proper lighting conditions which cannot be controlled outdoors. Therefore, two RGB streams from different angles at given intersection will be used and will be complemented with one thermal stream. The heat image from the thermal stream complements the RGB streams well during the night and furthermore sidesteps some of the privacy concerns that show up in systems like these.

In short, this project will study the usage of computer vision to do data reduction on traffic data with focus on vulnerable road users using thermal and RGB traffic scene data.

2. The scientific content of the PhD project

2.a. Background for the project problem

The project is about generating an automatic analysis of the vulnerable road users in the traffic scene. The project can be broken down into five stages:

1) Large scaled multi-modal and multi-view data collection in traffic intersections.

2) Big data annotation using semi-supervised classifier.

3) Vulnerable road users(VRU) classification in the traffic scene using deep learning using multi-modal and multi-view data.

4) Semi-automatic VRU traffic analysis.

5) Using UAVs to generate traffic usage reports in roundabouts and intersections.

The first stage addresses capturing data with multiple cameras and modalities in traffic intersections. In this stage equipment must be analysed with the scope of creating a robust capturing platform that can be deployed at intersections and capture data using batteries. A battery-powered recording platform has already been developed in cooperation with the Traffic Research Group at Aalborg University, but needs to be developed further such multiple cameras can be included.

The second stage deals with a major research issue of how to annotate large amount of data in a proper manor. By manually annotate a small part of the dataset, and use this to create an object classifier which are applied yet another small part of the dataset. This workflow are iterated until the entire dataset are annotated.

The third stage will include developing a robust vulnerable road users classification (VRU) framework. The classification module will be based on deep learning methods and the framework must be contain a graphical user interface such traffic researchers can use of it.

The fourth stage deals with developing of a semi-automatic VRU traffic analysis framework for traffic researchers. They will use it for data reduction on their large datasets, the framework must be able to adapt to changing environments by determining and adjusting parameters automatically.

The fifth and final stage will be to use UAV (Unmanned aerial vehicle), i.e. a drone, to capture data at roundabouts or intersections in RGB and thermal. This data should be used as input to an automatic analysis tool, which can generate traffic reports based on the data, e.g. queue length on each roundabout leg, traffic density, entrance and exit in the traffic scene, etc.

2.B. INTRODUCTION TO THE STATE-OF-THE-ART FOR THE PHD PROJECT

Detection and classification of traffic scene objects have been a research topic for a long period of time, but remains an active research area [21, 20]. The research area involves the topic of detecting and classifying objects, e.g. vehicles, both as a driving assistance system(DAS) and driving analysis study point of view. Methods from both topics are considered useful when surveying the state-of-the-art.

Detection is usually split into two broad approaches: appearances- and motion-based methods. Where some DAS motion-based applications make use of optical flow [9] to detect moving objects and Kalman filter for tracking them [1].
The appearance-based methods make use of appearance features such as edges[14], Histogram-of-Oriented-Gradients (HOG)[24, 9], Gabor[24], Haar-like[2, 19], or SURF[14]. The features are mostly used with Adaboost classification[2, 19] and Support Vector Machines (SVM)[24, 9] for detecting and classifying vehicles and other objects in the traffic scene. The HOG-SVM combination have in particular been very popular from a broad array of object detections, but other methods such as integral channel features(ICF) and subsequently the aggregated channel features(ACF) have been applied to detecting large variation of traffic scene objects, such as traffic signs[16], traffic signals[10], and pedestrians[6].

The above approaches have however for the past years been challenged and outperformed in most benchmarks by convolutional neural networks (CNN), e.g. PASCAL VOC 2012 [8]. The very prominent CNNs have been applied in a broad array of application, but mostly detection, classification, and semantic segmentation, where they outperform most hand-crafted features. In [3] semenatic segmentation is also done using DCNN and fully connected conditional random fields (CRF). The best performing methods on semantics classes: bicycles, bus, car, motor bike, and persons, in the PASCAL VOC 2012 utilizes deep convolutional neural networks (DCNN) and conditional random fields (CRF), which is implemented as a Recurrent Neural Network (RNN), [27]. CRFs have been very popular to combine with DCNNs, but other methods such as markov random fields (MRF)[15] has also seen been applied. The MRFs have in general been outperformed by the publications using CRFs in most of the semantic classes, but the current best performing on cyclists segmentation are using MRF[15] with 61.6 %.

The use of multi-modal deep learning for object recognition is a relatively new research area, where the two most common modalities are RGB camera and hardware acquiring depth, e.g. stereo or lidar, such as in [7, 26] for object recognition. In [17] both RGB, depth, and thermal are used for face recognition, which outperforms their previously published results.
But in general, the use of multi-modalities for traffic scene object classification has seen limited use compared to using solely RGB. In addition, the multi-view RGB settings has seen very limited use when fused together with thermal and compared to UAV data [12].

2.C. PROJECT'S OBJECTIVES

To investigate methods using multi-view RGB and thermal data, as well as data capturing from UAV and 360°cameras to classify traffic scene objects, such as cars, trucks, cyclists, and pedestrians. The final goal is to implement the investigated state-of-the-art methods into vulnerable road users classification (VRU) framework, such the latest computer vision state-of-the-art methods are made available for traffic researchers. The foundation of this work will rely on comprehensive data collection and experiments with different modalities and view angles.

The concrete research questions are:

- How can large amount of data be annotated in a semi-supervised manner?
- How can classification of traffic scene objects be improved using multi-modal and multi-view data to classify?
- How do object classification on UAV data perform compared to single-view and multi-modal and multi-view data?

2.D. KEY METHODS

The project will include state-of-the-art methods for traffic scene object classification such as CNNs and CRFs [27]. Multiple CNNs will be used based on multiple modalities, i.e. RGB and thermal, and view angles. The capturing setup will consists of two recordings platforms with 1 RGB and 1 Thermal cameras, and 1 RGB Camera, which will be mounted on existing infrastructure in intersection selected based on high conflict and accidents statistics. The results of using multiple CNN streams will be investigated based on both early feature fusion and late decision fusion [22].

From a traffic researcher point of view, most interesting conflicts and accidents occurs when multiple events are triggered in a specific time order. The above methods will therefore be applied to related traffic scene problems such as traffic light detection where recent work has relied on ACF[11]. Other related areas include cyclists, motor bike detection, and pedestrian detection, where a comparison between above methods will be compared to the HOG-SVM [4] and ACF [5] approaches.

The expected amount of data collected during this PhD project is enormous, which requires the development of an automatic image annotation (AIA) framework, which will be developed during this PhD project based on above the investigated methods e.g. CRFs[13] and DCNN[18]. The AIA framwork must be used for pixel-level annotation across modalities and views, such it can be used directly for training CNNS and CRFs for traffic scene object classification and semantic image segmentation.

Finally for all of the above methods, possible tracking methods, e.g. hidden markov models and Kalman filters, are expected to be used for reducing false positives as well as helping predicting and interpolating in the semi-supervised annotating framework.

2.E. POTENTIAL SIGNIFICANCE AND APPLICATION(S)

As mentioned in section 2.b, a lot of work have already been done in using a single RGB stream for vehicle classification. Few have however combined multiple RGB and thermal streams from different viewpoints in an intersection. Furthermore, acquiring traffic data from UAVs and comparing the results with traditional stationary image acquisition methods are an unexplored area.
Finally, a VRU toolbox will be developed throughout this thesis. This application will become open-source and is expected to have high significance as it provides an automatic traffic analysis tool for traffic researchers, which is a very limited application area. Thus, many sub-problems are expected to show up, giving plenty of opportunity for future research.

2.F. WORK AND TIME PLANS

The time plan for this PhD Project is divided into 6-months sections with associated milestones. Each period is described with a small introduction and bullets presenting the specific milestones.

MAY 2016 - OCTOBER 2016 Data collection, preliminary analysis for VRU classification, and finally prototype of trafic flow tool for roundabouts based on 360°recordings.

- Data collection for VRU classification.
- Perform interview with traffic researchers to determine possible future parameters.
- Preliminary analysis and experiment in cooperation with ATKI regarding traffic usage reports in roundabouts.
- Develop traffic flow analysis tool for ATKI regarding traffic usage reports in roundabouts.

NOVEMBER 2016 - APRIL 2017 Comparison of UAV data and stationary data as well as traffic light detection.

- Preliminary UAV data collection.
- Perform occlusion analysis of data collected from UAV and stationary.
- Automate heuristic parameters for VRU traffic analysis tool.
- Deep learning detector for traffic light detection.

MAY 2017 - OCTOBER 2017 Automatic image annotation and automation of Road User Behaviour Analysis tool.

- Creation of annotated multi-viewed traffic intersection dataset.
- Automate current heuristic parameters in Road User Behaviour Analysis tool.
- Deep learning for traffic scene object classification.

NOVEMBER 2017 - APRIL 2018 Detecting and classifying using one camera.

- Creation of annotated multi-viewed traffic intersection dataset.
- Perform annotation variance analysis.
- Develop traffic scene object classifier for big data.

MAY 2018 - OCTOBER 2018 Detecting and classifying using multiple modalities and views.

- Perform analysis and develop of cyclists segmentation.
- Develop traffic scene object classifier using multi-modal and multi-view streams.

NOVEMBER 2018 - MARCH 2019 Detect and high level queue analysis based on UAV recordings.

- Final UAV data collection.
- Develop automatic traffic usage report tool for UAV recordings.
- PhD Thesis writing.

APRIL 2019 PhD Thesis done.

For each of the milestones, 1-2 papers will be written and submitted to computer vision conferences or journals.

2.G. THESIS OUTLINE

This thesis will consist of articles written during the PhD period.
Possible outline:

- Introduction.
- Understanding Vulnerable Road Users: A Computer Vision Study.
- Annotating big data using semi-supervised classifier.
- Segmenation and classification of traffic scene objects using multi-view and multi-modal data.
- Using UAV data to automatically determine traffic flow.

2.H. TENTATIVE TITLES

Table 2.1: Overview of tentative titles with belonging abstract, responsible author(s) and title type. Titles marked with green are titles that are already published.

Title	Type	Author(s)	Abstract
Night-time drive analysis using stereo vision for data reduction in naturalistic driving studies	Paper	**Morten B. Jensen**, Mark P. Philipsen	In order to understand dangerous situations in the driving environment, naturalistic driving studies (NDS) are conducted by collecting and analyzing data from sensors looking inside and outside of the car. Manually processing the overwhelming amounts of data that are generated in such studies is very comprehensive. We propose a method for automatic data reduction for NDS based on stereo vision vehicle detection and tracking during day- and nighttime.
Vision for Looking at Traffic Lights: Issues, Survey, and Perspectives	Journal	**Morten B. Jensen**, Mark P. Philipsen	This paper presents the challenges that researchers must overcome in traffic light recognition (TLR) research and provides an overview of ongoing work. The aim is to elucidate which areas have been thoroughly researched and which have not, thereby uncovering opportunities for further improvement. An overview of the applied methods and noteworthy contributions from a wide range of recent papers is presented, along with the corresponding evaluation results. The evaluation of TLR systems is studied and discussed in depth, and we propose a common evaluation procedure, which will strengthen evaluation and ease comparison.
Aggregated Channel Feature Parameter Depletion on Traffic Lights during Nighttime	Paper	**Morten B. Jensen**	We have used the ACF detector in most of previous published work, but without tweaking all parameters individually. We wish to generate a heatmap with the optimum ACF parameters on the nighttime data from LISA Traffic Light Dataset.
Regional Convolutional Neural Networks and Conditional Random Fields on Traffic Light and signs during day- and nighttime	Paper	**Morten B. Jensen**	We have used the ACF detector in most of previous published work. In this paper we introduce the use of CNN and CRF for detection on both the LISA Trafic Light and Sign dataset.
A Comparative Detection Study: Regional Convolutional Neural Networks and Aggregated Channel Features on Traffic Light and signs during day- and nighttime	Journal	**Morten B. Jensen**	We have used the ACF detector in most of previous published work. This is a more extended study compared to *Aggregated Channel Feature Parameter Depletion on Traffic Lights during Nighttime*. We introduce and compare ACF with CNN on both the LISA Trafic Light and Sign dataset.
Using Deep Learning in Applications: Traffic Scene Object Classification in the Road User Behavior Analysis Tool	Paper	**Morten B. Jensen**	Trafik og veje application paper

Multimodal Deep learning for Robust RGB-T Object Recognition in the Traffic Scene	Paper	**Morten B. Jensen**	Technical paper on using deep learning for classification on traffic scene objects using both RGB and thermal data. The classification is done by two seperate CNN streams, one for each modality, which are combined in a late fusion step. A comparison of the individual CNN streams and the late fused classifier is done.
A Large-scale Multi-view RGB-T Traffic Scene Object Dataset	Paper	**Morten B. Jensen**	Introduce a new dataset in a paper. Could be combined with a survey and published as a journal.
Ongoing Work on Multimodalled Deep Learning for Object Recognition in the Traffic Scene	Paper	**Morten B. Jensen**	A small survey paper gathering important work done in using 2 or more modalities in deep learning for object recognition in the traffic scene.
Object Recognition in the Traffic Scene using Multi-view RGB and Thermal data	Journal	**Morten B. Jensen**	Using 1 thermal CNN stream and 2 RGB CNN streams from different positions in a traffic scene to do object recognition.
Multimodalled Deep Learning in Applications: Object recognition using Multi-view RGB and Thermal data	Paper	**Morten B. Jensen**	Trafik og veje application paper
UAV Versus Stationary Traffic Recordings: A Comparative Study	Paper	**Morten B. Jensen**	By collecting data captured from different viewpoints in an traffic intersection at the same time, one can make a comparative study of the detection and tracking performance. An interesting analysis is done based on adjusting the degrees of occlusion.
UAV Versus Stationary Traffic Recordings: A Comparative Study	Paper	**Morten B. Jensen**	Trafik og veje application paper

RELEVANT CONFERENCES: ICCV, CVPR, ECCV, AVSS, IV, ITSC, WACV, MVA, Trafik og veje.

RELEVANT JOURNALS: ITS, ISPRS, PAMI, CVIU, MVA.

3. SUPERVISOR/STUDENT CO-OPERATION AGREEMENTS

The main supervisor is Professor Thomas B. Moeslund.

The supervisor and the student will meet at the office when necessary during the daily work. At least once a month, a scheduled meeting will be held where the status of the project will be discussed. If the supervisor is out of office for more than a few days, the student will be informed, and likewise if the student is out of office.

The supervisor will suggest appropriate journals and conferences for publications of results, and will provide supervision and comments on papers.

4. PLAN FOR PHD COURSES

Table 4.1: Overview of PhD Courses

Course	**Place**	**ECTS**	**Type**	**Status**
Introduction to the PhD Study	AAU	1	General	Completed
Writing and Reviewing Scientific Papers	AAU	3.75	General	Planned
Professional communication	AAU	2.5	General	Planned
Management of Research and Development	AAU	2.5	General	Planned
Scientific Computing Using Python - 1. Python + Scientific Computing	AAU	2.5	General	In progress
Scientific Computing Using Python - 2. High Performance Computing in Python	AAU	2	General	In progress
Machine Learning	AAU	3	Project	Planned
Bayesian Statistics Simulation and Software	AAU	4	Project	Planned
Deep Learning Summer School	Montreal	5	Project	Planned
Internal VAP study group	AAU	3	Project	In progress
Workshops and conferences		3	Project	In progress
Total:		32.25		

Further information on the Deep Learning Summer School can be found at: `https://sites.google.com/site/deeplearningsummerschool2016/`

The internal VAP study group is a biweekly study group in the Visual Analysis of People Laboratory coordinated by Professor Thomas B. Moeslund, where members of the lab presents state-of-the-art papers, their current work, and other relevant topics.

5. Plan for fulfilment of knowledge dissemination

Project supervision, first as a co-supervisor, later as main supervisor. Supervision will take place in the early Medialogy semesters, Information Processing Systems (6th semester) or Vision, Graphics and Interactive Systems (Master's), and finally 4th semester Robotics. Furthermore TA on relevant courses (computer vision, image processing, programming), and finally presentations at various conferences. The results of this project is also to be reported in lay-man terms to the involved traffic analysts as well in reports to the European Commission.

6. Agreements on immaterial rights to patents

IPR is handled via the standard university rules.

7. External co-operation

As 2/3 of this PhD project is founded through an European project, cooperation with several external partners are necessary throughout the PhD project. This leads to the possibility of multiple stays abroad with the other European partners. Currently the most promising stays involves a stay with our European partner TNO in Netherlands and a stay with Professor Mohan M. Trivedi at UC San Diego, USA.

8. Financing Budget

The PhD project is founded 2/3 by the In-Depth understanding of accident causation for Vulnerable road users(InDeV) project and 1/3 internally by the AD:MT department. The InDeV project is funded by the EU Framework Programme for Research and Innovation, HORIZON 2020, in grant agreement number 635895.

References

[1] J. Diaz Alonso, E. Ros Vidal, A. Rotter, and M. Muhlenberg. "Lane-Change Decision Aid System Based on Motion-Driven Vehicle Tracking". In: *IEEE Transactions on Vehicular Technology* 57.5 (2008), pp. 2736–2746. ISSN: 0018-9545. DOI: `10.1109/TVT.2008.917220`.

[2] W. C. Chang and C. W. Cho. "Online Boosting for Vehicle Detection". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 40.3 (2010), pp. 892–902. ISSN: 1083-4419. DOI: `10.1109/TSMCB.2009.2032527`.

[3] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs". In: *arXiv preprint arXiv:1412.7062* (2016).

[4] Navneet Dalal and Bill Triggs. “Histograms of oriented gradients for human detection”. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. Vol. 1. IEEE. 2005, pp. 886–893.

[5] Piotr Dollár, Ron Appel, Serge Belongie, and Pietro Perona. “Fast feature pyramids for object detection”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.8 (2014), pp. 1532–1545.

[6] Piotr Dollár, Serge Belongie, and Pietro Perona. “The Fastest Pedestrian Detector in the West.” In: *BMVC* 2.3 (2010), p. 7.

[7] Andreas Eitel, Jost Tobias Springenberg, Luciano Spinello, Martin Riedmiller, and Wolfram Burgard. “Multimodal deep learning for robust rgb-d object recognition”. In: *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE. 2015, pp. 681–687.

[8] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. *The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results*. URL: `http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html`.

[9] Amirali Jazayeri, Hongyuan Cai, Jiang Yu Zheng, and Mihran Tuceryan. “Vehicle detection and tracking in car video based on motion model”. In: *IEEE Transactions on Intelligent Transportation Systems* 12.2 (2011), pp. 583–595.

[10] Morten B Jensen, Mark P Philipsen, Chris Bahnsen, Andreas Møgelmose, Thomas B Moeslund, and Mohan M Trivedi. “Traffic light detection at night: Comparison of a learning-based detector and three model-based detectors”. In: *International Symposium on Visual Computing*. Springer International Publishing. 2015, pp. 774–783.

[11] Morten Bornø Jensen, Mark Philip Philipsen, Andreas Møgelmose, Thomas B Moeslund, and Mohan M Trivedi. “Vision for Looking at Traffic Lights: Issues, Survey, and Perspectives”. In: *IEEE Transactions on Intelligent Transportation Systems* (2016).

[12] Konstantinos Kanistras, Goncalo Martins, Matthew J Rutherford, and Kimon P Valavanis. “Survey of unmanned aerial vehicles (uavs) for traffic monitoring”. In: *Handbook of Unmanned Aerial Vehicles*. Springer, 2015, pp. 2643–2666.

[13] Wei Li and Maosong Sun. “Semi-supervised Learning for Image Annotation Based on Conditional Random Fields”. In: *Image and Video Retrieval: 5th International Conference, CIVR 2006, Tempe, AZ, USA, July 13-15, 2006. Proceedings*. Ed. by Hari Sundaram, Milind Naphade, John R. Smith, and Yong Rui. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 463–472. ISBN: 978-3-540-36019-3. DOI: `10.1007/11788034_47`. URL: `http://dx.doi.org/10.1007/11788034_47`.

[14] Bin-Feng Lin, Yi-Ming Chan, Li-Chen Fu, Pei-Yung Hsiao, Li-An Chuang, Shin-Shinh Huang, and Min-Fang Lo. “Integrating Appearance and Edge Features for Sedan Vehicle Detection in the Blind-Spot Area”. In: *IEEE Transactions on Intelligent Transportation Systems* 13.2 (2012), pp. 737–747. ISSN: 1524-9050. DOI: `10.1109/TITS.2011.2182649`.

[15] Ziwei Liu, Xiaoxiao Li, Ping Luo, Chen Change Loy, and Xiaoou Tang. “Semantic Image Segmentation via Deep Parsing Network”. In: *CoRR* abs/1509.02634 (2015). URL: `http://arxiv.org/abs/1509.02634`.

[16] Andreas Møgelmose, Dongran Liu, and Mohan Manubhai Trivedi. “Detection of US traffic signs”. In: *IEEE Transactions on Intelligent Transportation Systems* 16.6 (2015), pp. 3116–3125.

[17] Marc Oliu Simon et al. “Improved RGB-D-T based Face Recognition”. In: *IET Biometrics* (2016). ISSN: 2047-4946.

[18] George Papandreou, Liang-Chieh Chen, Kevin Murphy, and Alan L. Yuille. “Weakly- and Semi-Supervised Learning of a DCNN for Semantic Image Segmentation”. In: *CoRR* abs/1502.02734 (2015). URL: `http://arxiv.org/abs/1502.02734`.

[19] S. Sivaraman and M. M. Trivedi. “A General Active-Learning Framework for On-Road Vehicle Recognition and Tracking”. In: *IEEE Transactions on Intelligent Transportation Systems* 11.2 (2010), pp. 267–276. ISSN: 1524-9050. DOI: `10.1109/TITS.2010.2040177`.

[20] Sayanan Sivaraman and Mohan M Trivedi. “A review of recent developments in vision-based vehicle detection.” In: *Intelligent Vehicles Symposium*. Citeseer. 2013, pp. 310–315.

[21] Sayanan Sivaraman and Mohan Manubhai Trivedi. “Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis”. In: *IEEE Transactions on Intelligent Transportation Systems* 14.4 (2013), pp. 1773–1795.

[22] Cees GM Snoek, Marcel Worring, and Arnold WM Smeulders. “Early versus late fusion in semantic video analysis”. In: *Proceedings of the 13th annual ACM international conference on Multimedia*. ACM. 2005, pp. 399–402.

[23] Danmarks Statistik. *Danskerne får flere og yngre biler*. URL: `http://www.dst.dk/da/Statistik/NytHtml?cid=26472` (visited on 06/14/2016).

[24] Zehang Sun, G. Bebis, and R. Miller. “Monocular precrash vehicle detection: features and classifiers”. In: *IEEE Transactions on Image Processing* 15.7 (2006), pp. 2019–2034. ISSN: 1057-7149. DOI: `10.1109/TIP.2006.877062`.

[25] trafiktypen.dk. *Cykel*. URL: `http://www.trafiktypen.dk/cykel` (visited on 06/15/2016).

[26] Ziyan Wang, Ruogu Lin, Jiwen Lu, Jianjiang Feng, and Jie Zhou. “Correlated and Individual Multi-Modal Deep Learning for RGB-D Object Recognition”. In: *CoRR* abs/1604.01655 (2016). URL: `http://arxiv.org/abs/1604.01655`.

[27] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr. “Conditional random fields as recurrent neural networks”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015, pp. 1529–1537.