Lifelong Learning for Dynamic Object Detection and Tracking for Autonomous Vehicles

Prof. Yassine Ruichek
EPAN Research Group, https://epan-utbm.github.io/
Université de technologie Belfort-Montbéliard (UTBM), France

Situational awareness is crucial for safe and reliable operation of autonomous vehicles in urban environments. In particular, detection and tracking of dynamic objects are the most important aspects to be considered, as the vehicles share its operational space with human actors in various forms, such as pedestrians, cyclists, and cars. Situational awareness is typically achieved through the perception and environment modeling system of the autonomous vehicles, which has to: 1) collect environmental data of the surrounding area using on-board sensors, 2) detect the other actors in the sensory data in a reliable and timely manner, and 3) track the actors to analyze and predict their motion. These three capabilities constitute the core competency of an autonomous vehicle that has to reliably operate in normal traffic situations.

However, different sensors exhibit different performances under different environmental conditions, and some of them might fail to provide relevant data in difficult, but realistic situations such as during adverse weather and dense traffic. Unreliable sensory information makes early detection and robust tracking of dynamic objects a difficult task.

To mitigate the disadvantages of individual sensors, autonomous vehicles combine heterogeneous sensors into multi-sensory systems. Such combination significantly enhances the perceptual abilities (*e.g.* robustness, range, accuracy) of autonomous vehicles. Unfortunately, the data of some of these sensors are not easy to interpret, because they do not provide complete information about the 3D positions of the perceived objects (*e.g.* cameras) or their data do not resemble information provided by human senses (*e.g.* 2D/3D LiDAR, ultra wide band radars). However, recent advances in machine learning methods proved methods capable of segmenting specific objects from sensory information and can be learned automatically, given that there are enough training data available [1].

In this thesis, we will investigate machine learning methods for multi-sensory systems for autonomous vehicles, where the heterogeneous nature of the sensory data will allow mutual training of the methods of dynamic object detection and tracking. On-the-fly, lifelong learning of the objects models for detection and tracking will be achieved through exploitation of the heterogeneity and amounts of data gathered by the vehicle's sensors over long periods of time. In contrast to existing technologies [1, 2, 3, 4] which are mainly based on static models, our research work will focus on development of adaptive models that are completed and refined based on the data gathered over long-time operation of the autonomous vehicle.

The thesis work will combine three different sensors installed in our autonomous car: 1) 3D LiDAR scanner, which provides direct 3D information about the object surfaces in the car vicinity, but has a limited resolution and range, 2) color camera, which provides high-resolution 2D data with unlimited range, but is affected by adverse weather conditions, and 3) ultra wide band radar, which provides data that are not affected by weather conditions, but are difficult to interpret.

In order to achieve lifelong mutual learning, we will develop an online learning framework, which updates objects models by exploiting dynamic objects trajectories fused from multi-sensor detections. On the one hand, at first, the model has to be initialized by supervised learning with labeled subjects (e.g. pedestrians, cyclists, cars). The initial learning set can be small though (e.g. only one sample), as more samples will be incrementally added and used for updating the models in future iterations, making lifelong learning possible. On the other hand, we need a robust multi-target tracking system to gather the detections, fuse them and generate their tracks. This system will play an important role in our solution framework, because it associates the detections across time as well as across the different sensors, making lifelong training and mutual sensor learning possible.

This PhD thesis is part of the Intelligent Vehicle Project investigated by LE2I-CNRS[1] laboratory at UTBM[2] (EPAN Research Group[3]). The project consists in developing advanced concepts and methods for driving assistance and automated driving. The research activity in this project is based on perception and localization using multi-source data collected from several sensors (camera, 2D/3D LiDARs, GPS, etc.) and knowledge systems like digital maps. The project is developed and implemented using several automated vehicles equipped with different sensors. The Figure below shows one of our experimental automated vehicles equipped with several sensors.

Supervisors:

- Prof. Yassine RUICHEK, Head of the EPAN Research Group
  - Email: yassine.ruichek@utbm.fr
  - Homepage: `http://epan.fr/People:Ruichek_yassine`

[1] `http://le2i.cnrs.fr/`
[2] `http://www.utbm.fr/`
[3] `https://epan-utbm.github.io/`

References

[1] P. Ondruska, J. Dequaire, D. Z. Wang, and I. Posner, "End-to-end tracking and semantic segmentation using recurrent neural networks," *arXiv preprint arXiv:1604.05091*, 2016.

[2] A. Teichman and S. Thrun, "Tracking-based semi-supervised learning," in *Proceedings of Robotics: Science and Systems*, 2011.

[3] A. Petrovskaya and S. Thrun, "Model based vehicle detection and tracking for autonomous urban driving," *Autonomous Robots*, vol. 26, no. 2-3, pp. 123–139, 2009.

[4] D. Z. Wang, I. Posner, and P. Newman, "Model-free detection and tracking of dynamic objects with 2d lidar," *International Journal of Robotics Research (IJRR)*, vol. 34, no. 7, pp. 1039–1063, 2015.

EPAN Research Group's related recent published papers

- Ibrahim Kajo, Nidal S. Kamel, **Yassine Ruichek**, Aamir Saeed Malik. "SVD-based tensor-completion technique for background initialization". *IEEE Transactions on Image Processing*, 27(6): 3114-3126, 2018.

- Abdelmalik Moujahid, Fadi Dornaika, **Yassine Ruichek**, and Karim Hammoudi. "Towards semantic segmentation of orthophoto images using graph-based community identification". *Neural Computing and Applications*, 2017.

- Citlalli Gámez Serna and **Yassine Ruichek**. "Dynamic speed adaptation for path tracking based on curvature information and speed limits". *Sensors*, 17(6):1383, 2017.

- **Zhi Yan**, Li Sun, Tom Duckett, and Nicola Bellotto. "Multisensor online transfer learning for 3D LIDAR-based human classification with a mobile robot". Submitted to *IROS 2018*, available on `https://arxiv.org/abs/1801.04137`.

- Li Sun, **Zhi Yan**, Anestis Zaganidis, Cheng Zhao, and Tom Duckett. "Recurrent-OctoMap: Learning state-based map refinement for long-term semantic mapping with 3D-lidar data". *IEEE Robotics and Automation Letters*, 2018.