# EARLY DETECTION OF CARS EXITING ROAD-SIDE PARKING

*Matossouwé Agninoube Tchalim, Sio-Song Ieng, Jean-Philippe Tarel*

COSYS-PICS-L
Univ Gustave Eiffel
F-77454 Marne-la-Vallée, France
Email: Jean-Philippe.Tarel@univ-eiffel.fr

## ABSTRACT

Vehicles suddenly exiting road-side parking constitute a hazardous situation for vehicle drivers as well as for Connected and Autonomous Vehicles (CAV). In order to improve the awareness of road users, we propose an original cooperative information system based on image processing to monitor vehicles parked on the road-side and on communication for sending early warning to vehicles on the road about vehicles leaving their parking space. We have implemented and tested this system in two places in France with parallel and perpendicular parking slots for several camera positions, and we report on its efficiency and its limits.

***Index Terms***— Vehicle detection and recognition, tracking, event detection, early detection of departing vehicles.

## 1. INTRODUCTION

Intersections and entries have been known as hazardous points of the road networks for over a century. A diversity of solutions have been developed to reduce the number of accidents at intersections, with traffic lights or roundabouts, and entrance ramps for motorway. Vehicles exiting suddenly from a street parking are more difficult to handle with infrastructure improvement and lead to hazardous situations for vehicle drivers as well as for Connected and Autonomous Vehicles (CAV). Camera-based road side systems able to improve the awareness of road users at road intersections and in roundabouts have already been proposed [1], but nothing has been proposed, to our knowledge, to improve awareness of vehicle leaving on a road-side parking slot.

Cameras were used for instance for an improved management of the entries, exits and occupancy of closed parkings [2], but not for road-side parking. In [3], an ultrasonic sensor attached to the side of the vehicle was used for occupancy statistics of road-side parking free slots.

Thanks to project ENA (Expérimentations de Navettes Autonomes) for funding. ENA Project is funded by the Government as part of the Future Investments Program, now integrated into France 2030, and operated by ADEME.

**Fig. 1**. The information system consists in a camera attached to a public lighting pole, watching road-side parking slots in the City of Biot, France. The camera is connected to a computer in a box and a road-side unit (added on the image).

The foremost objective of the proposed information systems is to improve driving safety by broadcasting information about vehicles exiting from road-side parking slots. This information collected by road users allows to improve the knowledge of their environment, and thus can be used to take safer and more efficient driving decisions. The proposed system may also be used for the purpose of traffic monitoring to collect occupancy statistics of road-side parking free slots.

In this paper, we focus on information systems dedicated to the detection of exits from road-side parking slots, based on one or several cameras linked with a communication unit able to broadcast early warnings to connected road users. In Sec. 2, the camera-based information system and video processing algorithms are described. Then, Sec. 3 describes the experiments performed in two places in the Cities of Biot and Champs-sur-Marne, France during the project ENA, *Expérimentations de Navettes Autonomes*, and the achieved efficiency and observed limits.

## 2. CAMERA-BASED ROAD-SIDE INFORMATION SYSTEM

We have designed a camera-based road-side information system able to detect vehicles exiting their road-side parking slots and able to broadcast an early warning to all mobiles with a receiver. The goal of this information system is to broadcast the accurate position of the vehicle exiting the road-side parking space and the time of detection.

As shown in Fig. 1, the road-side information system consists in a camera in a waterproof case attached to a pole at 6 meters, a computer connected to the camera for video acquisition and processing also in a waterproof box attached to the pole, a road-side unit for sending warnings to the road users with a receiver. The sooner the information about a vehicle exiting its parking space is broadcast, the more efficiently the road user can evaluate the risk and react accordingly. The time of detection is thus an important information which must be broadcast along with the parked vehicle position. The delay between detection and broadcast is a critical parameter of the system. But the main critical parameter which determines the usefulness of the system is the delay between the detection of a potential departure and the actual departure. Another critical parameter is the frequency at which the information about warnings is refreshed.

### 2.1. Camera acquisition

For completeness, several cameras should be installed depending of the size of the road-side parking and of the field of view the camera-based system is able to process correctly. We have investigated two camera positions: parallel or perpendicular to the road axis. The camera should be installed more than 6 meters above the ground to avoid occlusion by high trucks or buses. The camera lens should be of fixed focal length, and its field of view should be selected to cover no more than the parking slots of interest. We have experimented two kinds of road-side parking slots: parallel and perpendicular. The camera should be with automatic iris and gain tuning to adapt to the variety of lighting conditions. The camera can be connected to the computer with a IP cable up to 100 meters (possibly with a PoE injector) or with a USB cable up to 5 meters. Each image must be time stamped by the camera at the time of acquisition, unless the transmission is fast enough (within a few milliseconds) to allow a time stamp by the connected computer. The time on the camera and on the computer must be synchronized, using NTP or PTP for more accuracy. For broadcasting, the computer time must be synchronized with the GPS time.

### 2.2. Detection and recognition in images

The video captured by the camera is streamed to the computer to perform the first step of the processing. Parked vehicles and pedestrians have to be detected in every frames of the video. For this, we tested detection and recognition using the well-known state-of-the-art convolutional neural network (CNN) named YOLO (You Only Look Once) which has the advantage of speed [4]. YOLO predicts both a probability map of the presence of the objects of interest, several possible bounding boxes and possible classes. The learning thus consists in the optimization of the sum of the following loss functions: the class recognition loss, the boundary box loss and the presence of object loss.

We experimented with YOLOv5 version which is easy to interface within the processing chain. The detections in each frame are bounding boxes, each with a class label. We reduced the number of classes from 80 to five: pedestrian, car, bicycle, motorbike, truck/tractor/bus. We performed a refinement of the learning on our own labeled dataset to improve detection and recognition performance on selected places. Our labeled dataset was built from images extracted from a few videos from on site cameras. The images were selected to include various vehicles and various positions in the road-side parking slots.

### 2.3. Detection tracking along images

The detection being performed on each image, detected objects need to be tracked along frames in order to estimate speed. We used Deep SORT [5] for object tracking in successive frames because it is fast and reliable after fine tuning of the parameters to the camera view. Deep SORT consists in two steps: a prediction step using a linear Kalman filter followed by an association step between new detections and existing tracks using the Hungarian algorithm. The association step takes into account the geometric distance, the speed and the visual similarity from the YOLO features. Each newly detected object is assigned with a new tracking index and it keeps its index until it leaves the camera field of view. However, it may lose its index value if it is occluded for too long. The resulting outputs are tracks with bounding box, class label, tracking index and speed. For each track, the age of apparition can be computed as the number of frames from track beginning.

### 2.4. Early detection of vehicles exiting parking

To achieve early detection, rather than detecting a car starting, our idea is to detect when a pedestrian enters his car. Indeed, the delay between a driver entering his car and the car departing is higher than several seconds. As shown in Fig. 2, the status of a mobile, moving (status label 'Mobile') or standing still (status label 'Stop'), can be guessed from its track by setting a threshold on its estimated speed. When a pedestrian track finishes close to a stopped car track, the driver entry is detected. The closeness is obtained with a threshold on the IoU (Intersection over Union) between the car and pedestrian bounding boxes. The tracking age of the pedestrian should be large enough to consider the track, to avoid perturbation from

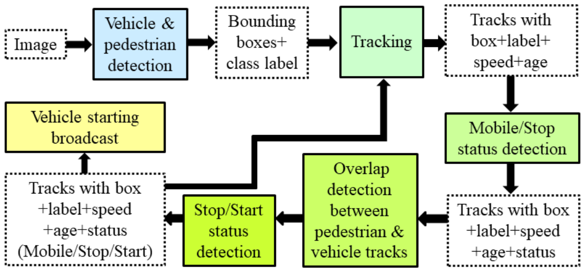


**Fig. 2**. The video processing consists in vehicle and pedestrian detection, followed by tracking. Then, tracks are processed to detect when a pedestrian is stepping inside his car.

index changes due to missed detection on a few frames. To remove false positive, pedestrians detected close enough to a bicycle or a motorbike are also discarded. After the detection of driver entry, the status of the car track is set to be starting, with label 'Start'. We also tested a variant where an extra constrain is added: the status of a pedestrian should be 'Stop' before they can be consider entering their car.

### 2.5. Intrinsic camera calibration

To be able to estimate the position of detected vehicles on the ground, the camera transformation between the ground and the image should be estimated. The intrinsic camera calibration consists in estimating the parameters of the camera model, such as the pixel size, image center and distortions. We used the pinhole camera model with three terms for radial distortion. Performing the intrinsic calibration involves taking snapshots of the same chessboard plane at various positions and orientations [6]. It is performed only once, before the camera is fixed inside its waterproof case.

### 2.6. Extrinsic camera calibration

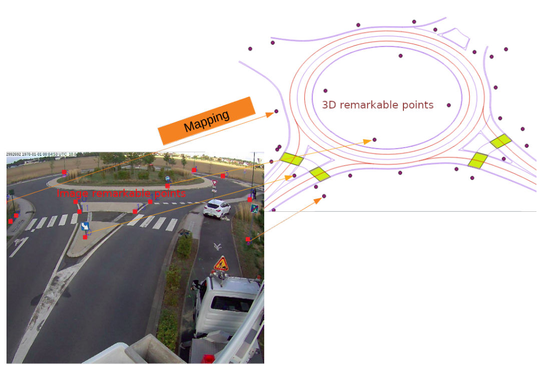


**Fig. 3**. The image coordinates of remarkable points are associated with their 3D coordinates to estimate the camera position and orientation.

The extrinsic camera calibration consists in estimating the position and the orientation of the camera with respect to a reference coordinate system on the ground. Therefore, it must be performed in situ every time the camera is moved.

To be able to provide useful pose information, a shared reference coordinate system should be used such as the one used in GPS. Performing the extrinsic camera calibration involves selecting remarkable points of the scene such as corners at of road signs, street lamps and corners of lane markings, in the camera image. As shown in Fig. 3, the image coordinates of these image points are associated to their 3D coordinates obtained for instance from a RTK GPS. From this set of paired 2D and 3D coordinates, knowing the intrinsic camera parameters, the position and orientation of the camera can be estimated by minimizing the error between the 2D image coordinates and the 3D coordinates projected into the image. Nelder-Mead's Simplex algorithm [7] was used for this optimization.

### 2.7. Message broadcast

When the status of a car track is with label 'Start', a message is broadcast with the car position, current time and time of driver entry. The size of the backprojected bounding box being related to the height of the detected object, we decided to use the backprojection of the mid point of the bottom segment of the bounding box as the detected position on the ground, a choice we found to give the most robust results. Messages are broadcast using a road-side unit (RSU) such as the ones from Lacroix City which can deal with information about 255 perceived objects using the Collective Perception Message (CPM) format, at the maximal frequency of 10 Hz.

## 3. EXPERIMENTS

Experiments were conducted, within the ENA project, along the Roumanille avenue in the city of Biot, with the help of Eiffage (see Fig.1), and also along the Newton boulevard in the city of Champs-sur-Marne. Only the acquisition and processing steps of the system were tested, but not communications.

The evaluation of early detection of cars exiting parking was performed on a 7.5-hour video acquired in one day in Champs-sur-Marne, after building the manual ground-truth. It was also performed on the 7.5-hour video acquired on may 9th 2022 in Biot. An example of early detection is shown in Fig. 4 in Champs-sur-Marne and in Fig. 5 in Biot, where cars with 'Mobile' status are in blue, cars with 'Stop' status are in magenta, cars with label 'Start' are in red and pedestrians are in green. In these two examples, the driver is detected entering the car 30 s and 20 s before the car exits its parking slot. We observed an average delay between driver entry and car exit of 52 s with a standard deviation of 32 s, from 20 observations. This delay is long enough to allow the processing

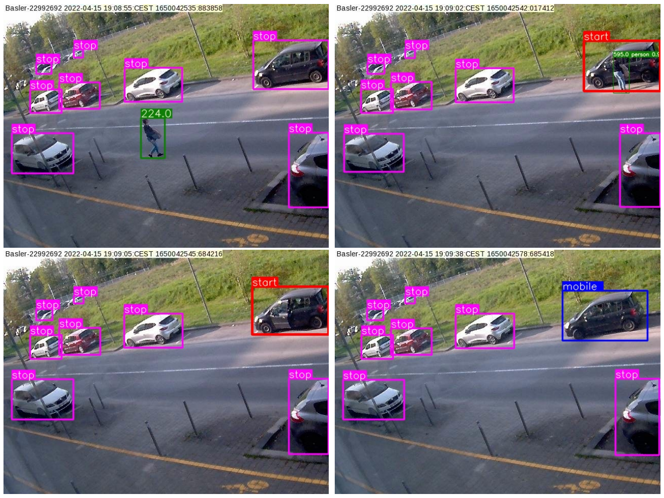


**Fig. 4**. Early detection (30 s before exit) of a car exit in Champs-sur-Marne, with parallel road-side parking slots.

and broadcast frequency to be reduced.


**Fig. 5**. Early detection (20 s before exit) of a car exit in Biot, in parallel and perpendicular road-side parking slots.

For Champs-sur-Marne, the achieved precision is 58% and 64% using the detection algorithm variant. In both cases the recall is 100%. All targets are detected with few false detections thanks to the camera being perpendicular to the road which allows to clearly observe entries in cars in a 30 m range of parallel road-side parking slots, see Fig. 4. When the camera is parallel to the road, cars are more occluded, especially for angle and non-parallel parking slots, see Fig. 5. This make it difficult to correctly detect each vehicle and pedestrians between the vehicles. The first variant is thus no longer effective and results degrade with distance. For Biot, the precision and recall are 28% and 95% on a distance range of 30 m, but 23% and 92% for a range of 60 m. Therefore, the system should be implemented with a camera perpendicular to the road.

## 4. CONCLUSION

We have described an original camera-based system able to broadcast early warnings about vehicle exiting road-side parking slots. The proposed system consists in a video camera, a computer for the video processing and a road-side unit able to broadcast messages. The different steps of the video processing have been described and discussed, in particular the capacity for early detection and low refreshing time. The proposed system was evaluated in terms of detection performance for parallel and perpendicular parking slots in the two cities of Biot and Champs-sur-Marne in France. We intend to evaluate the proposed system in degraded weather conditions and to study how it can be improved.

## 5. REFERENCES

[1] Sio-Song Ieng, Mathias Paget, Matossouwé Agninoube Tchalim, and Jean-Philippe Tarel, "Improved awareness for safer intersection of flow of mobiles," in *submited to 2023 IEEE Intelligent Vehicles Symposium (IV'23)*. IEEE, 2023.

[2] Khaoula Hassoune, Wafaa Dachry, Fouad Moutaouakkil, and Hicham Medromi, "Smart parking systems: A survey," in *2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA)*. IEEE, 2016, pp. 1–6.

[3] Suhas Mathur, Tong Jin, Nikhil Kasturirangan, Janani Chandrasekaran, Wenzhi Xue, Marco Gruteser, and Wade Trappe, "Parknet: drive-by sensing of road-side parking statistics," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, 2010, pp. 123–136.

[4] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015.

[5] Nicolai Wojke, Alex Bewley, and Dietrich Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*. IEEE, 2017, pp. 3645–3649.

[6] Zhengyou Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the seventh ieee international conference on computer vision*. Ieee, 1999, vol. 1, pp. 666–673.

[7] John A. Nelder and Roger Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.