Journal of Automation, Electronics and Electrical Engineering

Design of application for vehicle classification with queue detection using deep learning

Artur SKORUPKA, Eng. [1]; **Tomasz CIECHULSKI**, PhD Eng. [1*]

[1] Military University of Technology, Faculty of Electronics, gen. Sylwestra Kaliskiego 2, 00-908 Warsaw, Poland,
*tomasz.ciechulski@wat.edu.pl

DOI: https://doi.org/10.24136/jaeee.2025.011

Abstract – The aim of the study was to create a solution that would enable the detection of a queue of vehicles before an intersection with traffic lights in the Matlab programming environment. The work focuses on designing an application for detecting a queue of cars, using the YOLO algorithm and developing a mechanism for changing lights based on the detected queue. The solution presented in the paper can be applied in practice and contribute to increasing safety, traffic flow and efficiency of urban traffic management.

Key words – vehicle and queue detection, image classification, deep learning

Introduction

According to a report by the Polish Automotive Industry Association [1], over 26.3 million motor vehicles were registered in Poland in 2023. This is more than 10% of vehicles registered in the European Union. This number is constantly growing. There are almost 3.3 million cars in the Mazowieckie Voivodeship alone. Such a large number of vehicles contributes to an increase in traffic, traffic jams and reduced road safety. The challenges faced by designers of intersections and traffic management systems in cities are increasing. The use of modern technologies, such as vehicle detection at intersections, offers the opportunity to optimise urban traffic management. The development of deep learning algorithms and their application in image analysis enables effective detection and supports the automation of motion management processes. The motivation to take up the topic was the need to create a solution that would increase the safety and efficiency of intersections with traffic lights. The choice of detection technologies based on deep learning algorithms makes it possible to integrate such systems into smart city infrastructures, which is in line with the concept of modern security systems.

1. Queue detection methods

One of the most commonly used methods of vehicle detection is inductive detection [2, 3]. Induction loops are wires embedded in the surface in the form of a rectangle or other figure, through which alternating current flows, generating a magnetic field. A car passing through a magnetic field causes eddy currents, which leads to a change in the inductance of the loop. The most important advantages of the inductive loop detector include:

- insensitivity to weather conditions, which is related to the fact that the sensor is located a few centimetres under the road surface,
- relatively low price, until the detector needs to be repaired,
- high accuracy of measurements,
- simple design,
- simple signal conditioning circuits.

The biggest drawbacks of this method include the need to interfere with the road surface to place induction loops under the surface, and the lack of detection of people or bicycles, especially with a carbon frame. To detect bicycles, special induction loops are used, which are quite often installed under the surface of bicycle paths, especially in Denmark or the Netherlands. Such loops are properly calibrated to detect smaller vehicles such as bicycles, but still an induction loop will have a "problem" with a carbon frame bike. The phenomenon of eddy current (Foucault's current) is used to detect the vehicle. This phenomenon occurs when a metal object is located in the area of an alternating magnetic field. A magnetic field, encountering an object made of metal on its way, penetrates it, causing eddy currents to be induced. Eddy currents create a magnetic field that opposes the primary field and as a result weaken the original field, causing the sensor impedance to change both in terms of module and phase angle.

Another frequently used method of detecting vehicles at an intersection involves analysing the video stream from cameras placed above the intersection [4]. This method uses image processing and machine learning algorithms. Machine learning algorithms such as YOLO, SSD or Mask R-CNN are used, which allow for fast and precise detection of vehicles in real time [5]. This method allows to observe several lanes without the need to interfere with the structure of the surface. Another advantage of this detection method is the ability to analyse the speed of vehicles as well as the type structure of traffic. It allows to determine the size of the vehicle, the time the vehicle stays in a place or distinguish different types of vehicles, such as: passenger cars, trucks, buses, cyclists, pedestrians or emergency vehicles, and even read license plates. Thanks to such information, it is possible to determine what part of the total number are vehicles belonging to a specific group and, on this basis, better adapt the traffic light model to the specific requirements of a given intersection. This gives much greater opportunities for creating intelligent transport solutions. An additional advantage, that detection using inductive loops does not have, is the ability to recognize both public transport vehicles and emergency vehicles and the ability to prioritize them when changing the signalling phase. The biggest disadvantages include: sensitivity to weather conditions and hardware requirements. Fog, rain or too much sunlight can negatively affect image quality and make it difficult to detect vehicles, and real-time image processing requires a lot of computing power, especially when we have to analyse each frame of the recording to determine the traffic intensity. The situation with vehicle detection for excited signalling seems to be much easier. It allows to analyse frames even one second apart, which will make the calculations 30 times faster, and the delay of one second in the case of excited signalling should not cause problems for people waiting for the light to change.

Radio detection is most often used to detect trams. In contrast to induction loops, it allows a lot of information from the vehicle to be transmitted both to the traffic control room and to the local traffic light controllers. In addition, thanks to this method, it is possible to detect a public transport vehicle among other vehicles, for example in the case of a bus on a road without a designated special lane only for buses. In addition, it enables early detection of a vehicle approaching traffic lights without the need for additional infrastructure in the form of an induction loop under the track. The biggest disadvantage of radio detection is the need to equip vehicles with appropriate devices, such as autocomputers and radio antennas, which is associated with additional costs for each vehicle. This method is successfully used in public transport in cities in Poland (e.g. Kraków, Wrocław).

RFID detection [6] is mainly used in logistics, although this type of detection could be used in public transport vehicles. To use this type of detection, the vehicle should be equipped with one or more active or passive RFID (Radio-frequency identification) tags, and RFID sensors with antennas should be placed at intersections [7]. The main advantages of this type of detection are the cost of the appropriate markers. Passive markers cost a few dozen cents, active ones are more expensive, but they also work at much longer distances. The sensors are able to receive a signal from multiple tags at once, reducing the number of sensors needed. It is enough to place one sensor at one intersection, there is no need to install separate sensors for each of the inlets. An interesting concept is to place RFID tags in emergency vehicles, and antennas - in front of intersections, which would enable smooth passage through the intersection for emergency vehicles [8].

The method of detecting vehicles using optical radiation is based on visible or infrared radiation. The detector consists of a transmitter and a wave receiver, but they can be placed in one device. Detection with the use of optical radiation makes it possible to both detect a vehicle and determine the number of passing vehicles, their length or the presence of pedestrians. Such detectors provide higher measurement accuracy than inductive loops, detection errors for optical systems are less than 1% in good lighting conditions, detection errors for inductive loops are estimated at 2-5%. Optical radiation detectors are also resistant to electromagnetic interference. Unlike induction loops, they detect pedestrians or carbon bicycles, with which induction loops have a problem. The biggest disadvantage compared to induction loops is the higher cost of installation and further maintenance. Sensors must be cleaned and calibrated regularly to maintain their high accuracy. One of the systems that use optical radiation is LIDAR (Light Detection and Ranging) [9]. It uses laser beams to detect objects precisely. It allows to detect vehicles and determine their exact position and speed, detect pedestrians and create traffic maps in smart cities.

Gravity detection is a method of detecting vehicles based on recording changes in gravitational forces caused by the mass of the vehicle, in practice these systems are piezoelectric detectors, pneumatic quartz detectors or fibre optic detectors [10]. Pressure or strain gauge sensors record the weight of passing vehicles, so they can be used both to control traffic lights, monitor the occupancy of parking spaces and control the permissible weight of vehicles. Installation requires interference with the road surface, and intensive use can cause damage to the sensors, so such sensors are mainly used at small and medium-sized intersections, vehicle weighing stations, and parking lots. There is also a lack of advanced analysis of traffic parameters, such as the speed of vehicles or identification of public transport vehicles. The biggest advantages include the simplicity of the system, resistance to weather conditions and low maintenance costs.

Currently, they are mainly used in mass measurement systems, e.g. on the Millau viaduct in France.

2. VEHICLE DETECTION USING DEEP LEARNING

This chapter presents issues related to the development of an application for traffic light control. Initially the intersection was selected and the method of detecting the queue was chosen. The choice fell on the method of detection using cameras using artificial intelligence, so they needed an intersection from which recordings could be obtained from the perspective of a camera placed at such an intersection. A very good candidate for downloading the video material seemed to be the intersection of Rzymowskiego and Cybernetyki streets in Mokotów, Warsaw. Above it there is a pedestrian and bicycle bridge, allowing for convenient acquisition of the material from the right perspective (Fig. 1).

Fig. 1. View of Rzymowskiego Street, Warsaw from the perspective of the pedestrian and bicycle bridge

This material will be needed to simulate real traffic at the intersection, as it would be difficult to obtain permission to mount the camera at an existing intersection. The task of this application is to detect a queue of cars. The question seems obvious: what is a queue of cars? In general, a queue is defined as a set of entities waiting for something. In this case, we are talking about a collection of vehicles waiting for a green light. A collection of vehicles will be at least one vehicle. Therefore, we can only analyse the image when a red light is displayed at the entrance to the intersection.

This allows for analysis even every second, resulting in significant savings in computing resources. In another approach, the application analysed every few frames, while also checking the speed at which the cars were moving. Not moving or moving at low speed would mean that vehicles are waiting in line. The maximum speed of the vehicle in the queue was set at about 5 km/h, i.e. about 1.4 m/s. This meant that at 30 frames per second in the analysis of each frame, the vehicle travelled less than 5 cm, which would be an imperceptible movement. Processing every third or fifth frame was considered, which would give about 14 and 24 cm, respectively. Nevertheless, analysis every 30 frames and only when a red light is displayed is a better solution. In addition, we do not analyse the issue of vehicle movement, which saves even more computing resources. This leads to the fact that the application we will create will concern excited signals, because in this system we are not interested in the traffic intensity, but only in the presence of a queue, most often on a subordinate street. Although the selected intersection is suitable for obtaining material for analysis and simulation of the recording from a camera placed above the intersection, it is a large intersection - with 12 lanes and a large number of phases in the traffic light change cycle. At such an intersection, the acyclic signalling model, which requires a detailed analysis of traffic volume, would work best. The obtained material will therefore be used to analyse and simulate an intersection with less traffic, and the intersection model will be created on the basis of the intersection of Milenijna and Światowida streets in Warsaw, where the lack of traffic lights contributes to the occurrence of many dangerous situations.

The recordings were made at different times of the day and in different weather conditions. For the purposes of work on the application, it was possible to obtain about 6 hours of recordings for analysis. The recordings were made in Full-HD image resolution, at a speed of 30 frames per second. Initially, it was planned to use methods based on image analysis using image processing algorithms, such as edge detection, frame difference, or optical flow, which are functions built into Matlab. After analysing the pros and cons (Table 1), it was decided to use the YOLO (You Only Look Once) v3 model [11], built into the latest versions of Matlab [12], with the use of the so-called fine-tuning.

Table 1. Comparison of the pros and cons of image analysis methods

Method	Disadvantages	Advantages
Image analysis based on image processing algorithms	• sensitivity to weather and lighting conditions, • inability to classify objects, • limited use in ITS systems	• low computational requirements, • high efficiency in good weather conditions
Deep learning methods	• high computing requirements, • sensitivity to bad weather conditions (fog, snow)	• precision and versatility, • automatic classification, • possibility of use in ITS systems

It was decided to use methods based on deep learning, mainly due to the possibility of implementation in ITS (Intelligent transportation system) systems, which makes the application an interesting project in terms of further development and use. The YOLO v3 version was chosen because it is the latest version that is available in the Matlab program. Its main advantage is the balance of accuracy and speed, allowing to work with fewer computing resources. The Faster R-CNN (Faster Region-Based Convolutional Neural Network) model, which is a more accurate model, works much slower and requires a lot of computing power. The SSD (Single Shot Multibox Detector) model is less accurate than the YOLO models. YOLO models are widely used in real-time object detection due to their revolutionary approach with single image processing. They are used in autonomous vehicles, in sports analysis or in medicine to detect cancer [11]. The YOLO v3 model pre-trained on the COCO (Common Objects in COntext) dataset contains 80 categories of detected objects such as car, truck, bus or motorbike, which will be necessary in this application. In order to improve the detection of objects, the so-called fine-tuning is performed, i.e. training the model on the data we are interested in. To perform fine-tuning, a script was prepared that was used to divide the videos into frames every second at the resolution used by the standard YOLO v3, i.e. 416x416. Of the 21,600 frames obtained in this way, about 70%

were those that were repeated or there were no vehicles on them, so they were removed. After the final cleanup, about 5200 photos remained in the data, which were then divided into 4200 (80.8%) photos that were then to be used to mark the cars on them, the remaining 1000 photos (19.2%) remained for testing the model. The next step in preparing the data was to mark the positions of the vehicles in the photos with rectangular labels. For this purpose, the ImageLabeler tool has been opened in Matlab, thanks to which it can easily mark labels on photos. New labels were created in accordance with the YOLO model pre-trained on COCO, such as "car", "bus" or "truck", and manual labeling of the labels needed to train the selected model was started (Fig. 2).

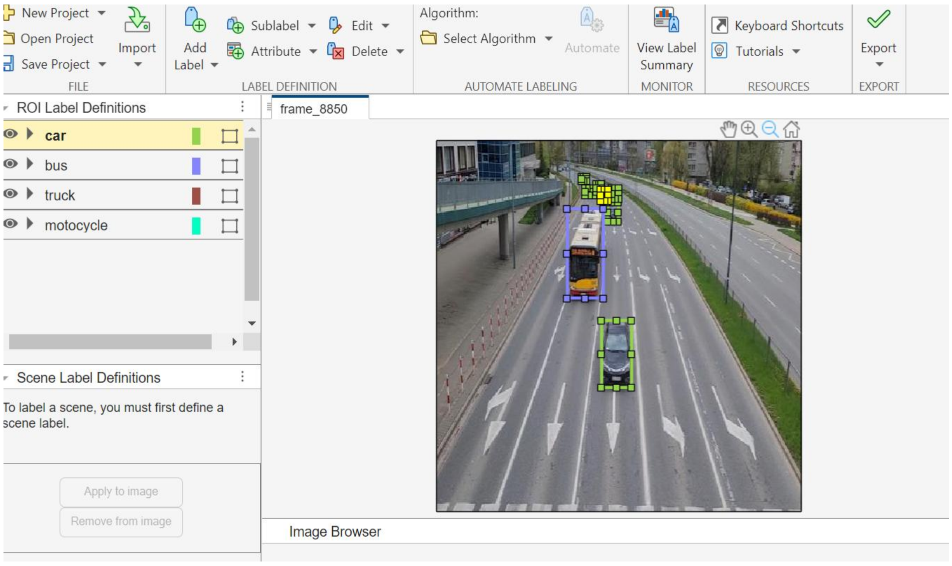

Fig. 2. An example of using ImageLabeler in Matlab

Before the model was trained, tests were carried out on the pre-trained model. For this purpose, cages were separated from the test set, covering the area from the vehicle stop line to the end of the continuous lines separating the lanes. This area covers about 20 meters from the traffic light. This is where cars stop to wait for a green light, so a queue forms. Then the tests began, the result of which is illustrated in Figure 3. Vehicles in all images from the test set were detected with an accuracy close to 100%, which resulted in the decision to abandon further model training and use the model pre-trained on COCO [13, 14] for the rest of the work. Model training is a time-consuming and computationally intensive process, and a pre-trained model can handle vehicles that take up a large part of the image.

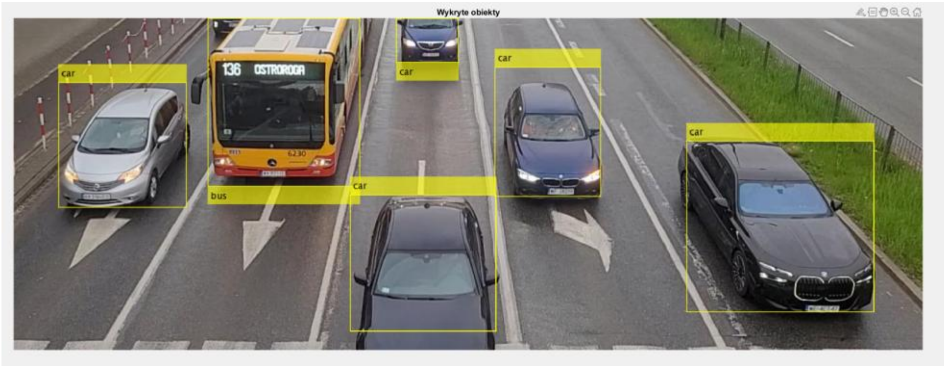

Fig. 3. Vehicle detection and classification with the YOLO v3 trained model

Therefore, if a small fragment of the recording is analysed, it should be effective enough to detect vehicles in one lane in the first 30 meters from the intersection. The YOLO model offers the possibility of further training if its effectiveness turns out to be insufficient, so it gives the opportunity to develop the application.

To test and present the results, several scenarios cut out from longer films were prepared, in which at the beginning there are no vehicles in the lane of interest, and they appear after some time and stop before the intersection so as to simulate a real situation in front of the traffic light. The scenarios are prepared at different times of the day and in different weather conditions, so as to faithfully recreate the conditions in front of the intersection.

3. RESULTS OF TESTING A YOLO MODEL PRE-TRAINED ON THE COCO SET

Testing the detection performance of the YOLO model was carried out on available data. An analysis of 300 images taken randomly from the test images was performed. The YOLO model was very good at detecting objects at a short distance from the intersection, i.e. those that took up a large part of the image. The model performed worse against objects located far away. Many vehicles occupying a smaller part of the image were often omitted from detection. As the queue forms in the immediate vicinity of the traffic lights, we are not interested in cars located at a considerable distance, i.e. outside the zone marked by solid lines. A vehicle there can change lanes in accordance with the regulations, so we do not even know which lane it will be in when it approaches the intersection. Therefore, the test area for model detection performance tests was set to about 30 m from the intersection, so that this area is larger than the final area in which the detection is carried out. The detection area in which the model performance test was performed is marked with a red line in Fig. 4.

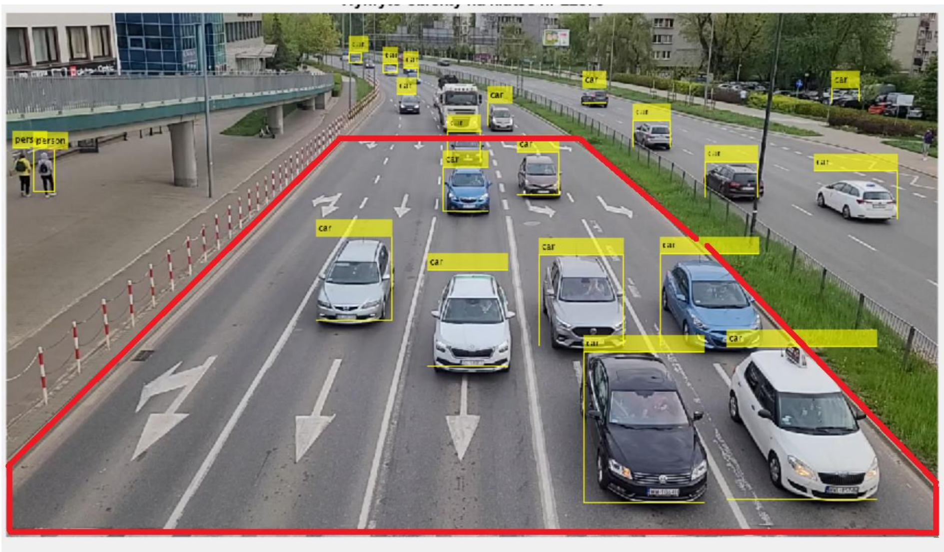

Fig. 4. Detection area marked with a red line

As can be seen in the example illustrated in Fig. 4, a truck that was at a considerable distance from the intersection was not detected, but all 9 vehicles within the zone of interest were detected correctly. In the tests, 1000 images taken at different times and in different weather conditions were analyzed. A total of 6534 vehicles were in the area of interest, which were both stationary and waiting for passage, as well as moving vehicles. The detection results were divided into 3 categories:

- correct detections,
- incorrect detections that do not critically affect the operation of the application,
- incorrect detections that critically affect the operation of the application and no detections.

There was 6438 situations which were correctly detected, which is 98.53% of all detections. Next 82 labels were incorrectly assigned, which is 1.26% of all detections. The most common error of this type was the failure to detect a vehicle obscured by another vehicle – 22 detections, which is 0.34% of all detections. Such an error can only happen when there is already a car on the cage, so it does not critically affect the operation of the application. Another mistake was the detection of a delivery vehicle up to 3.5 tons as a bus - 21 times and accordingly 0.32% of detections. Such an error does not cause a big problem, both the car and bus objects are in the set of vehicles that are detected, the only difference is in the assigned weight and such an error can only cause the light to change faster for the delivery truck. Four times the bus was detected as a train, which can be solved by adding an additional train class with a weight of 3 later in the development of the application. 82 errors are errors of this type, i.e. those that do not have a critical impact on the operation of the application. The remaining 14 bugs (0.21%) are nine detections of a motorcyclist as persons and this bug could already make the application not detect some motorcyclists. However, these errors only occurred in the upper part of the area that was analyzed, when the motorcyclist approached the intersection, the model detected both a person and a motorbike. 5 errors are errors where there was a complete lack of detection of the vehicle, once in the case of a very unusual truck/crane, another 3 cases when the vehicles were at a relatively large distance from the intersection, but already in the detection area. The case of the ladder on the roof of a car is particularly interesting. Such a simple action as placing a ladder on the roof of a car makes it difficult for the model to see from a distance, which can have many interesting applications. In all cases, the vehicles were correctly identified when they were directly in front of the traffic lights. There was no erroneous detection of vehicles in its absence, which means that all shadows, drops or birds flying in front of the camera do not interfere with the detection. Testing results are provided in Table 2.

Table 2. YOLO model testing results depending on lighting conditions

Weather conditions	Valid detections [%]	Mislabeled [%]*	Fatal error [%]**
Sunny South	98,97	0,81	0,22
Night	95,55	4,26	0,19
Rain in the morning	98,58	1,21	0,21
Averaged values for the entire set	98,53	1,26	0,21

Among all vehicles that appeared in the detection area, both correct detections (98.53% of cases) and incorrect label assignment were obtained, which does not critically affect the operation of the application, and lack of detection and assignment of an incorrect label, which critically affects the operation of the application defined as a critical error.
*The mislabelling group included detections:
- covered vehicle, 22 times (0.34%),
- vans as trucks, 21 times (0.32%),
- of a van as a bus, 20 times (0.31%),
- double detection of the car, 13 times (0.20%),
- bus as a train, 4 times (0.06%),
- truck as a car, 2 times (0.03%).

** The group of critical error, i.e. one that critically affects the operation of the application, included the following detections:
- motorcyclist as a pedestrian, 9 times (0.14%),
- no detection, 5 times (0.08%).

The best results were obtained in good weather conditions, with good sunlight, the worst results were at night, but these results were not critical from the point of view of the application. They would make the time until the lights change. As many as thirteen times, of which as many as 9 times at night (0.20% of all vehicles and 1.74% of vehicles at night), the model detected 2 cars in a situation when there was actually only one in that place. This type of error occurred mainly in night conditions. This error happened at different distances from the traffic lights, so it was not dependent on the distance, but it is another error that only causes the light to change faster than it should.

4. Conclusions

The use of queue detection on a large intersection in the left-turn lane can increase the capacity of the intersection in a situation where there are no vehicles on the left-turn lane. The green light time for the oncoming lanes can be extended by the green light time for the left turn lane and by the "between" times. In the situation with a secondary road, the lack of vehicles will also make the light for the direction with the priority road shine longer. The use of traffic lights in places where there are no such signalling will increase road safety without significantly reducing the capacity of the intersection. In the case of shuttle traffic, a certain change in the way the application works would have to be made. Checking what light is on the opposite signal, in the case of red light, the light change would take place without delay. In the case of green light, the light would change with a delay depending on the distance between the signals and the permissible speed.

The proposed method of queue detection using artificial intelligence algorithms enables both further work on the effectiveness of detection and implementation of the application into ITS systems. The YOLO model allows for a detection accuracy of 98.5%. Additional experiments showed that this result can be even improved by using fine-tuning. Additional model training on custom data could provide improved detection performance. According to the 2024 INRIX report [15], Warsaw is ranked 20th in the world, taking into account the average time spent in traffic jams, and it is getting worse every year. This time has increased by 15% since 2023 and by 20% since 2022. Traffic light models based on the analysis of data provided by detectors such as cameras are developed in parallel with other models, such as the INRIX Signal Analytics model, which uses anonymous data from vehicles connected to the system. The best strategy for the future will be to combine the advantages of several systems into one larger traffic management system.

Bibliography

[1] PZPM – Automotive Industry Report – https://www.pzpm.org.pl/pl/content/download/11094/68148/file/RAPORT_PZPM_07_12_2024_lekki.pdf access: 24.06.2025

[2] Lamas-Seco J.J., Castro P.M., Dapena A., Vazquez-Araujo F.J., "SiDIVS: Simple Detection of Inductive Vehicle Signatures with a Multiplex Resonant Sensor", Sensors, 2016, No. 16, 1309, https://doi.org/10.3390/s16081309

[3] Duda K., Marszałek Z., "Vehicle speed determination with inductive-loop technology and fast and accurate fractional time delay estimation by DFT", Metrology and Measurement Systems, 2024, vol. 31, no. 4, pp. 781-796, https://doi.org/10.24425/mms.2024.152048

[4] Mishra A., Chen K., Poddar S., Posadas E., Rangarajan A., Ranka S., "Using Video Analytics to Improve Traffic Intersection Safety and Performance", Vehicles 2022, 4, pp. 1288–1313, https://doi.org/10.3390/vehicles4040068

[5] Ventura R., Roussou S., Ziakopoulos A., Barabino B., Yannis G., "Using computer vision and street-level videos for pedestrian-vehicle tracking and behaviour analysis", Transportation Research Interdisciplinary Perspectives, 2025, 30, 101366, https://doi.org/10.1016/j.trip.2025.101366

[6] Paszkiewicz A., Pawłowicz B., Trybus B., Salach M., "Traffic Intersection Lane Control Using Radio Frequency Identification and 5G Communication", Energies, 2021, 14, 8066, https://doi.org/10.3390/en14238066
[7] Aleksandrowicz J., Piwowarczyk M., "Sposoby detekcji pojazdów transportu zbiorowego i ich funkcjonalność", Transport miejski i regionalny, no. 05/2016
[8] Neumann T., "Analysis of application possibilities of RFID technology in road transport", Scientific Journal of Gdynia Maritime University, no. 102/2017, pp. 44-60
[9] Ansariyar A., Taherpour A., "Statistical analysis of vehicle-vehicle conflicts with a LIDAR sensor in a signalized intersection", Advances in Transportation Studies, 60 (2023), pp. 87-106, DOI: 10.53136/97912218074246
[10] Szmigiel P., Szmigiel A., Stawowy M., "Wybrane metody identyfikacji pojazdów w systemie telematyki transportu", Logistyka, 2010
[11] Terven J., Córdova-Esparza D.-M., Romero-González J.-A., "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS", Machine Learning & Knowledge Extraction, 2023, 5, pp. 1680–1716, https://doi.org/10.3390/make5040083
[12] MathWorks, "MATLAB manual user's guide", Natick, 2024
[13] MathWorks, Getting Started with YOLO v3 https://www.mathworks.com/help/vision/ug/getting-started-with-yolo-v3.html access: 25.06.2025
[14] MathWorks, Object Detection Using YOLO v3 Deep Learning https://www.mathworks.com/help/vision/ug/object-detection-using-yolo-v3-deep-learning.html access: 25.06.2025
[15] https://inrix.com/scorecard/#city-ranking-list access: 30.06.2025