Algorithms for perception and detection of objects in self-driving electric vehicles

Aleksandar Stjepanovic, Miroslav Kostadinovic, Goran Kuzmic, Dejan Danilovic, Mario Radonjic
University of East Sarajevo
Faculty of Transport and Traffic Engineering
Doboj,BiH
aleksandar.stjepanovic@sf.ues.rs.ba,
miroslav.kostadinovic@sf.ues.rs.ba,
goran.kuzmic@sf.ues.rs.ba
danilovic.dejan@gmail.com
marioradonjic87@gmail.com

Marko Stjepanovic
University of Novi Sad
Faculty of Technical Sciences
Novi Sad,Serbia
stjepanovic.marko16@gmail.com

Abstract— **In order for vehicles to perform this function independently, it is necessary to implement complex systems that monitor the environment and make decisions independently. An important segment of this process are object detection algorithms. The subject of the research is the algorithms used in data analysis and decision-making during autonomous driving, as well as the application of these systems in electric vehicles. Algorithm components consist of a sensing component, that is, extracting significant information from raw sensor data; component of perception, which has the function of localizing the vehicle and understanding the current environment and components of decisions, in other words taking actions in order to reliably and safely reach the target destinations. Sensing component - usually an autonomous car consists of several main sensors.**

Keywords-autonomous systems, electric vehicles, sensors, neural networks, machine learning, algorithm models for autonomous driving

I. INTRODUCTION

Environmental protection is one of the main advantages of electric vehicles compared to fossil fuel vehicles. The overall costs of using electric vehicles are lower than the usage costs of traditional vehicles with regards to the mileage. The application of electric drives in vehicles, usage of artificial intelligence technology for implementation of self-driving smart vehicles, and the application of renewable energy sources, represent effective measures and adequate approaches in order to achieve sustainable transport of people and goods.

The development of intelligent electric vehicles can provide users with safety, comfort, and economy of vehicles with a significant reduction in environmental pollution. Furthermore, the capability of the autonomous driving system is useful in solving urban traffic congestion and useful for combining with the intelligent transportation environment of the future urban transportation system.

Autonomous intelligent vehicles are vehicles with technologies built into the vehicle that partially or fully enable the autonomous (independent) movement of the vehicle while increasing the safety of its use. Basically, autonomous intelligent vehicles are closely related to mobile robot technologies, and their use is usually based on the following four stages: Environment perception and modeling; Positioning and mapping with the help of GPS (Global Positioning System) technology; Vehicle route (trajectory) movement planning; and Movement control.

Autonomous vehicles use the Sense-Plan-Act algorithm that emerged with the development of Mobile Robot technology. This algorithm can be classified as AI (Artificial Intelligence) because it makes decisions and carries out activities without human participation.

For a vehicle to be autonomous, it must be able to "see - sense", "think", "drive" and when necessary, "learn", correct and improve its response. The "sense" of the environment is based on real-time data collection from on-board sensors such as visible spectrum cameras, microwave radar, LiDAR (Light Detection and Ranging), ultrasonic sensors, as well as precise mapping, tracking and other data from the environment streamed from the "cloud" service, such as road conditions and traffic information. Thanks to the development of vehicle communication technology V2X (Vehicle to everything), vehicles will be able to connect with other vehicles V2V (Vehicle to Vehicle), infrastructure V2I (Vehicle to Infrastructure), smartphones V2D (Vehicle to Device).

II. RELATED WORKS

In the paper [1], the authors provide an overview of existing perceptive methods that are used in the realization of autonomous electric vehicles and are used for the detection of 3D objects. The paper also analyzed and classified the mentioned methods with reference to future directions of research related to this area.

The authors in the paper [2] provide an overview of driving control systems and algorithms for smart electric vehicles, including advanced driving assistant systems, sensor implementation, vehicle dynamics, and control algorithms. The main contribution of this paper is reflected in the identification of trends to help researchers to define regulations in this area.

In the paper [3], the authors conducted research in the direction of autonomous driving, where an increasing number of publicly available datasets allow experiments with perception verification or for training and testing algorithms based on machine learning, while virtual test environments allow complete simulations of autonomous driving.

In the review paper [4], the authors analyze in detail the theory that is the basis of autonomous driving, the possibilities of implementing machine learning methods in object detection applications, and their classification, especially the classification of used data libraries, which plays an important role in the detection process.

Self-driving electric vehicles represent a synthesis of technologies that, in addition to a positive environmental impact, provide users with several advantages and benefits. In the paper [5], the authors analyze the trends in the development of modern means of transportation, where modern technologies for vehicle propulsion are included in synergy with Artificial Intelligence. The authors provide an overview of the state of technology available on the market and compare existing solutions with newly proposed ones.

In the paper [6], the authors analyze reliable Lane Keeping as one of the most important tasks for an automated vehicle. The authors' intent was to design and evaluate the performance of several different lane-keeping algorithms for an automated vehicle using computer vision. The performance of the implemented algorithms was evaluated and compared with tests performed with a human driver. The software packages ROS (Robot Operating System) and OpenCV were used for the software simulation of lane detection and keeping on the road. A test vehicle with a high-definition camera was used to implement and evaluate driving data. Each algorithm was evaluated based on several different parameters.

In the paper [7], the authors analyze the different needs of modern users, who use different types of micro-mobility (electric scooters) for short distances. For longer distances, modern transport services are used, where young users usually use services such as Uber and others by calling on a mobile phone. However, the distances between them are a few kilometers to about 10 kilometers and are not very practical for the use of the mentioned services and resources. The paper analyzes small electric vehicles with a shorter range and relatively low speed.

A startup project "Perceptln" dealing with a similar topic was also analyzed.

III. METHODS

The autonomous driving is not just one technology, but a very complex system consisting of several subsystems. Due to its complexity, it is convenient to divide it into three main components [8]:

- algorithms, including sensing, perception and decision making (requiring reasoning for complex cases);
- client systems, including the operating system and hardware platform;
- cloud platform, including high definition (HD) map, deep learning model training, simulation, and data storage.

The algorithm subsystem extracts meaningful information from the raw data it receives from sensors to understand its environment and make decisions about its future actions. Client systems integrate these algorithms together to fulfill requests in real time with an appropriate degree of reliability.

The algorithm component consists of sensors/senses, i.e. extracting significant information from the raw data coming from the sensors, then perception, which serves to localize the vehicle and understand the current environment, and finally decision-making, in other words, taking actions to arrive reliably and safely to the target destinations. In autonomous vehicles, data from multiple sensors must be combined to increase the reliability and safety of the decisions made [9].

Very important elements of the system for autonomous driving are the modules for predicting and directing the vehicle. The prediction module is responsible for predicting the behavior of identified surrounding objects in the perception module. The routing module performs lane-based routing and segmentation on high-resolution maps. Routing simply tells the autonomous vehicle how to get to its destination by following a series of lanes on HD maps. Its output is also used as input data for the Planning and Control modules. Figure 1 shows the modules for planning and control.

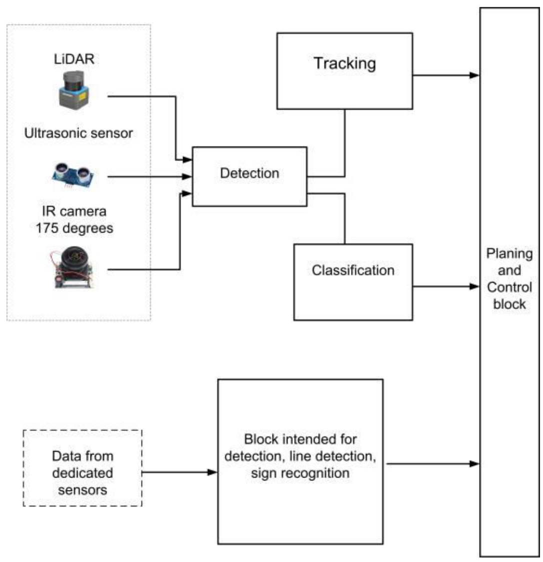

Figure 1. Planning and Control Moduls

The Perception subsystem (data processing infrastructure) consists mainly of three main blocks: detection, classification and tracking. Detection is the first step and deals with the separation of various objects from the environmental data provided by the sensors. For sensors such as cameras or LiDAR, the data is a matrix containing information for each point in the area covered by that sensor.

Sensors for Light Detection and Ranging (LiDAR) uses eye-safe laser beams to "see" the world in a 3D environment, providing machines and computers with a precise representation of the environment under investigation [10]. As one of the most important elements in the concept of modern autonomous vehicles, the LiDAR system is of great importance for the positioning of the vehicle and the environment in which the vehicle moves.

The working principle of the LiDAR system is based on measuring the time that elapses while the light passes from the transmitter to the emitted surface and back. Since the speed of light is much higher than the speed of sound, the time measured is very short, because light travels very fast, at speed of 300,000 km/s [11]. The sensor transmitter sends automatic rapid pulses of laser light to the surface at up to 150,000 pulses per second, the sensor on the instrument measures the amount of time it takes for each pulse to bounce back. The light travels at a constant speed so that the LiDAR instrument can calculate the distance between itself and the target with high precision [12].

By repeating this procedure at an extremely high speed, the instrument creates a complex map of the surface it has measured and thus creates a 3D image of the terrain. The interaction of the vehicle and all objects in its environment is an essential element for realizing the safe concept of autonomous driving. Sensor data is forwarded to the Perception phase to ensure the vehicle's comprehension of its condition and position, as well as the events surrounding the vehicle. The three main tasks in the Perception of autonomous driving are Localization, Object detection and Object tracking [10].

LiDAR sensors, which provide very accurate information about the state of surrounding space in depth, were initially mainly used for the detection and tracking of objects in autonomous vehicles. In recent years, with the rapid development of machine learning technologies, significant progress has been achieved in the process of object detection and tracking. Convolutional Neural Network CNN (Convolutional Neural Network) is a type of deep neural networks that are widely used in object recognition tasks. The general principle of CNN evaluation usually consists of the following layers [9]:

- The Convolutional layer contains different filters to extract different features from the input image. Each filter contains a set of "learnable" parameters,
- The Activation layer decides whether to activate the target neuron or not,
- The Data collection layer reduces the spatial size of the representation to reduce the total number of parameters to an acceptable number, thus enabling smooth operation of the network.
- Fully Connected layer where neurons have complete connections with all activation functions from the previous layer.

There are several main parameters for the tracking and operation of an autonomous vehicle, namely: the target vehicle distance R, the relative speed between the subject vehicle and the target vehicle V, and the target position of the vehicle defined by coordinates. These parameters must be used in safety control systems and can be obtained in signal acquisition and processing procedures. The frequency modulated signal, i.e. the transmitted signal, is reflected from the target vehicles. The target return signal is multiplied by a copy of the transmitted signal at the receiver. The frequency difference between the transmitted signal and the target return signals can be obtained using a mixer. This difference in frequencies is called the beat frequency. Figure 2 shows the architecture of the perception system [8, 12].

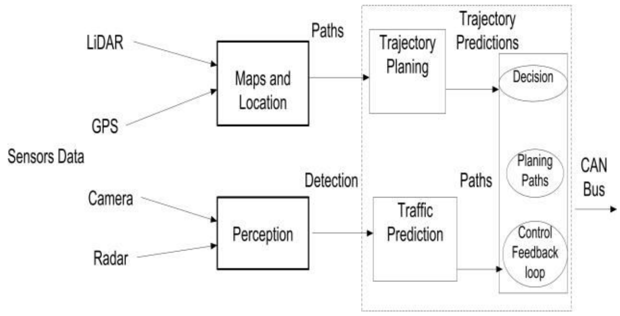

Figure 2. Perception system arhitecture [8]

The flow of data generated by an autonomous vehicle is of great importance. Classic vehicles that are tested today use many LiDAR, radar and camera sensors. These vehicles generate between 1 - 4 TB (terabyte) of data during one test. Based on the understanding of the vehicle's environment, the decision phase can create a safe and effective action plan in real time [12]. One of the main challenges when it comes to human drivers is to adequately respond to all the challenges that the driving process puts before them.

In the case of highly automated and fully automated driving, it is necessary for the vehicle to have the ability to recognize its own limitations enabled by its own machine perception, as well as the functional limitations of the data processing modules obtained based on this perception, and to react adequately based on the conclusions reached. Studies dealing with the analysis of highly automated driving simulators show that the real transition times to the driving performed by the driver are between 5 and 10s. In case of functional limitations, the vehicle would have to be able to reach an intrinsically safe state completely autonomously. However, potential transfer times of 5s and more require a very high autonomy of the vehicle, even if only in a small limited time.

Computer "Vision" is the identification and classification of objects through high-resolution computer cameras. This component consists of three phases. For the identification and classification of objects, the algorithm called "You Only Look Once" (YOLO) is most often used [13]. This algorithm divides the images into smaller square samples and each of these samples is passed through and analyzed by means of a Computational Neural Network (CNN) [14].

At this stage, all marginal square samples are present and none of the possibilities are excluded. The goal is to find the most accurate boundary pattern for each object. This can be done using a method called Non-Maximum Suppression (NMS). After implementing the YOLO and NMS method, the required result is obtained [14].

Object detection algorithms are broadly classified into two categories based on how many times the same input pattern (image) has passed through the network. Single-pass algorithms include the YOLO algorithm and SSD (Single-shot Detector).

On the other hand, the second group is represented byalgorithms that, in order to obtain object detection results, need a double pass of the image sample through the neural network. This group includes the following algorithms:

- R-CNN (Region-Convolutional Neural Networks)
- Fast R-CNN (Fast Region-Convolutional Neural Networks)
- Faster R-CNN (Faster Region-Convolutional Neural Networks)
- RF-CNN (Region-based fully-Convolutional Neural Networks)
- Mask R-CNN (Mask Region-Convolutional Neural Networks)

A characteristics of the R-CNN detector is to first generate sample regions using an algorithm such as Edge Boxes. Suggested regions are cropped from the image and resized. Next, the CNN network classifies the cropped and changed regions. Finally, the resulting frameworks for limiting the region proposals received some improvements using a Support Vector Machine (SVM) that was previously trained using CNN functions.

The Fast R-CNN algorithm also uses the Edge Boxes algorithm to propose sample regions. Unlike the R-CNN algorithm, which performs cropping and resizing regions, the Fast R-CNN algorithm processes the entire image. This algorithm is much faster than the classic R-CNN algorithm as its name suggests. The detector called Faster R-CNN generates through the RPN Regional Proposal Network, a proposal for the layout of region samples directly in the network, thus avoiding the use of some external algorithms such as Edge Boxes, which causes the algorithm to work even faster.

The algorithm called Mask R-CNN is a type of deep learning algorithm model, which is an extension of the Faster R-CNN algorithm and performs object detection and segmentation. Mask R-CNN combines object detection and segmentation into one, providing the ability not only to detect objects but also to precisely delineate their boundaries at the pixel level. Using FPN (Feature Pyramid Network) and OIAlign (Region of Interest Align), Mask R-CNN achieves great performance and accuracy.

Mask R-CNN has certain limitations, such as the complexity of the algorithm and the requirement for a large amount of memory during training and decision making. Also, during the operation of the algorithm, it is possible to encounter problems in the precise segmentation of very small objects or handling of heavily congested scenes. Acquiring large amounts of unsorted training data can be very demanding, and fine-tuning models for specific domains may require careful tuning of certain parameters.

IV. Discussion

Detection of objects in autonomous vehicles is crucial for safety and driving efficiency. Considering the precision and speed of detection, along with the characteristics of different algorithms, contributes to optimal vehicle management. The process of object detection in diverse scenarios requires complex algorithms capable of recognizing objects of various shapes, sizes, and textures.

This includes pedestrians, vehicles, cyclists, and other environmental elements. Algorithms based on deep learning (e.g., YOLO, SSD, Faster R-CNN) have high accuracy but can be resource-intensive, requiring powerful computing resources. Optimizing the execution speed of these algorithms is crucial for real-time driving. The combination of different sensors, such as cameras, LiDAR, and RADAR, is often used to obtain comprehensive information about the surroundings. Integrating this technology contributes to detection accuracy.

Different algorithms have their strengths and weaknesses. For example, deep learning can provide high accuracy, but model training requires time. Combining classical methods with deep learning can offer a balance between speed and accuracy. Vehicles face different driving conditions, including changes in lighting, weather, and road types. Detection algorithms must adapt to these changes to maintain a high level of accuracy.

Object detection is just one part of the vehicle control system. Integrating detection into the overall control system, including decision-making and action execution, is crucial for efficient driving. This process is directly related to driving safety, and errors in detection can have serious consequences. Therefore, it is necessary to develop algorithms that are resilient to errors and capable of recognizing critical situations. The combination of carefully chosen algorithms, sensor systems, and optimization for execution speed is crucial for successful object detection in autonomous vehicles. Continuous improvement in this field is necessary to achieve the highest possible accuracy and speed. Research in the field of self-driving cars is expected to evolve in several directions to address challenges and unlock new opportunities. Some key directions for future research in self-driving cars: Advanced Perception Systems, Safety and Emergency Handling, Regulatory and Legal Frameworks, Cybersecurity, Urban Mobility Solutions, Simulation and Testing.

V. CONCLUSION

The paper provides an overview of the main algorithms used in autonomous vehicles, that relate to perception or detection of objects. Also presented is a model of the architecture of the autonomous driving system perception module. Each of the mentioned algorithms has certain advantages and disadvantages.

Two groups of algorithms for object detection were most often used: with one pass through the network, where YOLO and SSD detectors are dominant, and the other group, which includes algorithms with two passes through the network, where R-CNN, fast R-CNN, faster R-CNN, region-based RF-CNN and Mask R-CNN. The mentioned region-based algorithms, starting with R-CNN, have certain limitations, especially regarding the speed of testing, so that R-CNN and Fast R-CNN algorithms are not suitable for real-time detection. On the other hand, the Faster R-CNN algorithm is characterized by high execution speed and is therefore suitable for detecting objects in real time.

Another group of algorithms such as YOLO and SSD algorithms have a different approach compared to region-based algorithms. These algorithms are characterized by a high speed of execution, but there are problems with the detection of small objects within the image (for example, a flock of birds).

REFERENCES

[1] D. Dai, Z. Chen, P. Bao and J. Wang, “A Review of 3D Object Detection for Autonomous Driving of Electric Vehicles”, MDPI, World Electric Vehicle Journal, doi.org/10.3390/wevj12030139, 2021.

[2] T. S. Haque, Md. H. Rahman, Md. R. Islam, Md. A. Razzak, F. R. Badal, Md. H. Ahamed, Ripon, S. I. Moyeen, Sajal K. Das, Md. F. Ali, Z. Tasneem, D. K. Saha, K. Chakrabortty, M. Ryan, “A Review on Driving Control Issues for Smart Electric Vehicles”, IEEE Access, Volume 9, 2021, pp. 135440 – 135472.

[3] Y. Kang; H. Yin; C. Berger, “Test Your Self-Driving Algorithm: An Overview of Publicly Available Driving Datasets and Virtual Testing Environments”, IEEE Transactions on Intelligent Vehicles, Volume: 4, Issue: 2, June 2019, pp. 171-185.

[4] A. Gupta, A. Anpalagan, L. Guan, A. S. Khwaja K. Elissa, “Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues”, Elsevier, Array 10, doi:10.1016/j.array.2021.100057, 2021.

[5] Ajao, Q., & Sadeeq, L., “Overview Analysis of Recent Developments on Self-Driving Electric Vehicles”. ArXiv. /abs/2307.00016, 2023.

[6] R. Kaddis, E. Stading, A. Bhuptani, H. Song, C. -J. Chung and J. Siegel, "Developing, Analyzing, and Evaluating Self-Drive Algorithms Using Electric Vehicles on a Test Course", IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS), Denver, CO, USA, doi: 10.1109/MASS56207.2022.00101,pp.687-692.

[7] L and J. -L. Gaudiot, "Autonomous vehicles lite self-driving technologies should start small, go slow," in IEEE Spectrum, vol. 57, no. 3, March 2020, doi: 10.1109/MSPEC.2020.9014458, pp. 36-49.

[8] S. Liu, L. Li, J. Tang, S. Wu, and J.-L. Gaudiot, “Creating Autonomous Vehicle Systems,”, A Publication in the Morgan & Claypool Publishers series, SYNTHESIS LECTURES ON COMPUTER SCIENCE, #12, 2020.

[9] M. Maurer, B. Lenz, J. Christian Gerdes, H. Winner, “Autonomous Driving,- Technical, Legal and Social Aspects”, Springer-Verlag GmbH Berlin Heidelberg, ISBN 978-3-662-48847-8, 2015.

[10] P. Wang, “Research on Comparison of LiDAR and Camera in Autonomous Driving,”, Journal of Physics: Conference Series, MAEIE 2021, doi:10.1088/1742-6596/2093/1/012032.

[11] L. C. Grinvald, Ofer T.-Sinai, “Smart Cars, Telematics and Repair Smart Cars”, University of Michigan Journal of Law Reform,Volume 54, 2021.

[12] J. Fraden, “Handbook ofModernSensors- Physics, Designs, and Applications”, Springer New York Heidelberg Dordrecht London, ISBN 978-1-4419-6466-3, 2010.

[13] P. Jiang, D. Ergu, F. Liu, Y. Cai, B. Ma, “A Review of Yolo Algorithm Developments”, Procedia Computer Science, Volume 199, 2022, Pages 1066-1073.

[14] Y. Zhou, S. Wen, D. Wang, J. Meng, J. Mu, R. Irampaye, “MobileYOLO: Real-Time Object Detection Algorithm in Autonomous Driving Scenarios”, MDPI, Sensors, 2022.