A Context-Aware Object Detection Method for Self-Driving Vehicles

Basha Alsahli and Olaa Motwalli

January 3, 2023

A Context-Aware Object Detection Method for Self-Driving Vehicles

*Basha Falah Alsahli[1], and Olaa Motwalli [1]
[1]Department of Computing and Informatics, University: Saudi Electronic University, Riyadh, Kingdom of Saudi Arabia
[1*]b.alsahli@hotmail.com and [1]o.motwalli@seu.edu.sa
* Corresponding Author

Abstract. Some issues with today's mobility are road accidents, fuel inefficiency, traffic congestion, etc., which directly or indirectly affect our economies and lives. AVs, also known as self-driving vehicles, can address these issues effectively. AVs take leverage from sophisticated sensor technologies. They classified perception systems for AVs into two classes: driving environment perception and positioning perception systems. In this paper, we addressed the driving environment perception problem. AVs need to drive safely through the roads. To avoid collisions, they need to identify various objects around them accurately. Therefore, we need a method that detects these objects with higher accuracy. Recent crashes of Tesla, Toyota, and Google self-driving cars indicate that a lot is required in order to make object detection methods for Avs. Hence there is massive scope for improvement. In the proposed work, we addressed object detection for AVs. The proposed object detection is a multiclass image segmentation problem. We used deep learning-based methods. The proposed method can be divided into (1) identifying context and (2) using context-based models for object detection. For performance evaluation, we used accuracy percentage, sensitivity, and specificity. The proposed method showed promising results at par with other schemes. The average prediction accuracy for ten classical deep-learning image frames is 89.021%. The average prediction accuracy for ten image frames of the proposed context-based deep learning model is 95.344%. We can see that the proposed context-based deep learning model produced 6.323% better accuracy than the base scheme.

Keywords: Autonomous Driving, Deep Learning, Object Detection, Context awareness, driving scene understanding, h2o.

1 Introduction

A critical function of autonomous vehicles (AVs) or self-driving cars is to accurately perceive their environment, including lane detection, detection of the navigable path, and recognition of static and dynamic objects, such as vehicles, bikes, and humans. Some issues with today's mobility are road accidents, fuel inefficiency, traffic congestion, etc., which directly or indirectly affect our economies and lives. AVs, also known as self-driving vehicles, can address these issues, effectively maintaining the highest passenger safety, efficiency, reliability, and sustainability. AVs take leverage from sophisticated sensor technologies [1]. Today the acceptability of AVs has increased among the public, particularly in developed economies around the world which is a positive sign if we have to see the large-scale success of AVs in the coming times [2].

A significant problem that needs to be handled on which the success of autonomous driving depends is that AVs' rides must be free from collisions and accidents. AVs need to drive safely through the roads. To avoid crashes, they first need to identify various objects around them with utmost certainty. These objects can be (other vehicles, obstacles, traffic lights, pedestrians, cyclists, driving spaces, etc.). Therefore, we need a method that detects these objects with higher accuracy. Recent crashes of Tesla [3], Toyota, and Google [4], Toyota [5] self-driving cars point out that a lot needs to be done to make object detection methods for AVs. Hence there is massive scope for improvement, which must be exploited.

Sensors have a critical function in AVs. With their help, AVs can sense the driving environment. This sensed data is further processed by intelligent algorithms whose task is to understand various contexts of the driving scene and assist autopilot in making a driving decision. These decisions can include reducing speed, turning, reversing, putting, and releasing brakes. To make the right driving decisions, two things are needed that are (1) accurate data sensing and (2) accurate driving scene prediction algorithms. Without quality data sensing, these machine learning-based algorithms do not have much worth. Cameras are the most fundamental sensor that is installed on AVs. It is an onboard sensor. Cameras are of various types, such as 2D, 3D, infrared, etc. The task of these cameras is to record visual data in the form of videos. Today various types of sensors are used in AVs to sense multiple data types, as depicted in Figure 1. AVs' most common sensors are Cameras, LIDAR, Radar, GPS, IMU, etc. Light Detection and Ranging (LIDAR) is a key but costly onboard sensor. It is the core of Google's self-driving car. LIDAR uses pulsed laser light to detect objects. Cost, short eyesight, and reliability are issues, as their accuracy is questionable in extreme conditions such as heavy rainfall, snowfall, storms, etc. However, today LIDAR is a much more powerful sensor than it was initially. Radar is a sensor that is preferred by Tesla self-driving cars over LIDAR. It is not expensive as LIDAR and has better, more extended eyesight and accuracy in extreme driving conditions. Radar uses radio waves to detect objects. However, detection accuracy at the corner is not as good as Lidar's. Other popular choices of sensors in AVs are the Inertial measurement unit (IMU), Global Positioning System (GPS), and Odometer. IMU sensors can measure several critical parameters related to angular rate, velocity, force, and acceleration and play a vital role in ensuring safe driving. GPS utilizes satellite data (longitude, latitude, speed, and direction) to navigate. The functions of IMU and GPS are fused to provide a safe driving experience.

Further Odometer is used to measure the speed of the AV. See (Y. Li & Ibanez-Guzman, 2020), (Dickmann et al., 2016), (Ertugrul & Ulkir, 2020), (Rahiman & Zainal, 2013), (Kutila et al., 2015), (Haltakov et al., 2012) for detailed explanation on various types of sensors used in AVs. AVs today are equipped with several sensors such as cameras, Lidar, Radar, GPS, etc. Particularly for driving scene understanding, AVs are using a fusion of cameras, Lidar, and Radar to make autonomous driving technologies. This multi-sensor requirement of AVs has a few drawbacks. Firstly, it increases the cost of the vehicles. Secondly, it increases the system's complexity. Knight (Knight, 2015) discusses how the technology industry now focuses on designing power single sensor perception technologies. In a demonstration showing how quickly some of the technology is advancing, Magna, a company that supplies components to most large carmakers, recently indicated that it can make a car drive itself on the highway using just a single camera embedded in its windshield. Magna hasn't said how much the technology would cost carmakers, but vehicle camera systems tend to cost hundreds of dollars rather than thousands. In Figure 1, we depicted various types of technologies needed for the proper function of AVs. Perception technologies are the most sensitive and vital technologies we still have to get proficient in. Further, we can classify perception systems for AVs into two classes: driving environment perception and positioning perception systems (J. Zhao et al., 2018). In this work, we will address the driving environment perception problem.

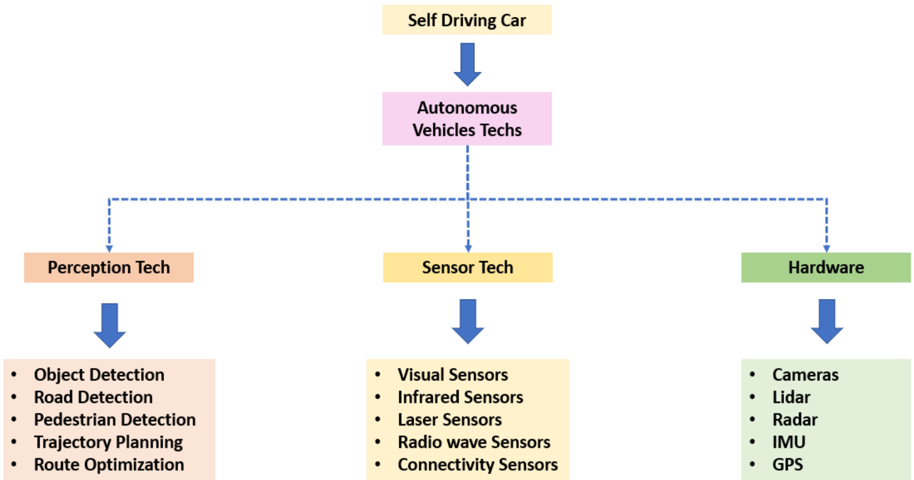


Figure 1. Various types of technologies needed for autonomous vehicles.

The rationale behind the proposed work is that if we want to replace human-driven vehicles with AVs, and driverless vehicles, we need to make the autonomous driving experience safer and hassle-free.

In the proposed work, we tried to address the problem of making autonomous driving collision-free and smooth. The primary challenge to this problem is to correctly identify various kinds of objects in a driving scene with high accuracy. To address this problem, we consider that context-awareness can help enhance object detection accuracy for AVs.

1.1 Contributions and Hypothesis

The contribution of the proposed work are:

- To critically review the existing literature concerning object detection for AVs to make the reader understand the current state-of-the-art.

- To propose a novel context-aware object detection method to produce higher prediction accuracy for safe and smooth driving.

- To design lightweight object detection methods through context awareness.

The hypothesis or the research question related to the proposed method is whether we can improve prediction accuracy by using a context-aware object detection method. The proposed object detection is a multiclass image segmentation problem and can be divided into phases. Firstly, we will download the dataset from the KITTI dataset repository for AVs. The next step is to label the data. Further, we will divide the dataset into training and testing. Then we will use machine and deep learning-based methods. One important thing about using a context-based model is that we do not need to train deep learning models on massive datasets. Instead, we will only train the model on more minor (context-specific data). Thus, our model will be lightweight because it will take less training and loading time. The brief details of the two phases are given below and later explained comprehensively in the methodology sections.

- In Phase-I, we will identify the context. By context, we are driving conditions, i.e., morning, night, extreme sunlight, shadow, etc., using a machine learning algorithm.

- In Phase II, once we have identified the context, we will use a context-specific lightweight deep learning model to detect various objects in the driving scene.

The paper is divided into five sections. In Section 2, we reviewed the kinds of literature on related topics. In Section 3, we will comprehensively explain the methodology of the proposed work. Whereas in Section 4, we will critically discuss the results and findings. Finally, in Section 5, we will give concluding remarks about the proposed work, its importance, its effects, and future recommendations.

2 Literature Review

Autonomous driving will play a significant role in the future, especially when we talk about high-level automation in smart cities, where most of our daily things will be automated [6]. However, today we need intelligent methods to perceive the driving environment with utmost certainty. Some challenging problems are object detection, trajectory prediction, collision avoidance, and traffic congestion avoidance are significant issues related to autonomous driving. The proposed work will address the critical object detection problem for autonomous driving. The initial problem is identifying various objects in the driving scene, known as object detection.

An object can be road, lane, pedestrians, cyclists, obstacles, vehicles, traffic signs, trees, animals, etc. The problem of object detection is challenging as the number of objects grows in the driving scene, making it harder for the classifier to predict accurately. Arnold et al. [7]. They performed a comprehensive survey on 2D and 3D object detection methods for autonomous vehicles. They review various machine learning techniques and their practical use in this field. An exciting work by Haris and Glowacz compared different object detection methods for AVs has been done in [8].

Machine learning and deep learning are at the forefront of detecting diverse objects in a driving scene. Classifiers such as Random Forest, Support Vector Machine (SVM), Decision Trees, and Ensembles are widely used for this problem [9]. Uçar et al. proposed an object detection method for Avs, a hybridization of CNN and SVM-based [10]. Further feature extraction is used to enhance the accuracy of prediction. One major challenge for autonomous driving is to drive safely in extreme weather conditions such as storms, snowfall, rain, sand storm, fog, etc. The accuracy of object detection systems decreases. To address this challenge, Walambe et al. [5] proposed an Ensemble-based machine-learning method for object detection in a driving scene to perform object detection in extreme weather conditions. Similarly, Masmoudi et al. [11] also proposed an SVM-based objection method for severe weather conditions.

Today Deep Learning algorithms are breaking all prediction records. They can understand complex and nonlinear data better than conventional machine learning algorithms. Due to these qualities, today, used in every field, from autonomous driving [12] to healthcare [13]. Fujiyoshi et al. [14] proposed a deep learning-based object recognition for autonomous driving using Convolutional Neural Network (CNN). Duhayyim et al. [15] presented robust deep learning (DL)-enabled object detection and classification (RDL-ODC) method with a particular focus on occluded or truncated objects.

Further, Park et al. [16] proposed a CNN-based object detection method that used fused data from cameras and Lidar. This work also tells about a new direction where researchers use data fusion techniques for object detection. Road and lane detection is a critical part of autonomous driving and can be seen as road detection's sub-problems. It becomes even more challenging in the absence of location information and on unstructured roads. We have witnessed machine learning algorithms used to detect lanes in recent years, but they failed to produce high efficiency and accuracy. Alam et al. [17] proposed TAAWUN, a novel road detection method that uses decision fusion and problem-specific feature selection based on deep learning and information sharing among nearby AVs. Vehicle detection is one of the most fundamental problems for AVs.

Another very challenging problem is Pedestrian and Cyclist detection, one of the most complex problems for a computer scientist in autonomous driving. The major problem, as identified by Ahmed et al. [18], with pedestrians and cyclists is that dynamic and sudden trajectory changes are too common and complex to predict due to their dependency on the intent of the individual pedestrian and cyclists. Further, cyclists are very thin objects, often looking like pedestrians and obstacles. Wang and Zhou [19] proposed a fast CNN-based pedestrian detection method using Beijing's urban traffic dataset. At the same time, Annapareddy et al. [20] used infrared cameras rather than typical cameras to map heat signatures can detect pedestrians, even in scenarios that are hard to identify. Ahmed et al. [21] used a novel data fusion method to detect pedestrians and cyclists based on Deep Neural Networks (DNNs).

Demers et al. [22] a moving vehicle in full-motion video. The researchers focused on increasing the chances of tracking and detection even in congested urban areas. They combined detection outputs from different spectral bands and related features to reduce false alarms. The authors used a GMM model for background pixel identification, which was then used for vehicle identification. The authors combined the components extracted from each spectral band to form a multi-spectral target region. The discovered target candidates were linked to targets from a tracking database by matching relationships between constructs from scale-invariant component changes.
Aytekin et al. [23] suggested investigating the importance of monochrome images collected with a single camera for vehicle detection and tracking. They focused on locating and tracking the vehicle during daylight hours, using data collected from within the car. They used vehicle shadow signals and related edge information to identify in practice quickly. They also included a lane detection model to address the issue of incorrect detection. After the road lanes are detected, the presence of a vehicle within the road area is estimated using the "shadow" as an indication. The authors used vertical edges to verify their approximate vehicle placement. The Kalman filter was used to monitor the detected features (from the vehicle area).
To extract the most moving objects, Kiratiratanapruk et al. [24] developed a background subtraction model using edge information. Boundary information was more resistant to illumination variations and required fewer computational resources than intensity-based background models. This allowed him to develop a method that could be used in real-time. Zhang et al. [25] suggested a cascade classifier ensemble-based classification model that is both robust and efficient. The first ensemble used in their model was heterogeneous, consisting of various classifiers such as kNN, Multiple-Layer Perceptrons (MLPs), SVM, and random forest. The classification was improved further in the second classifier ensemble, which consists of a collection of base MLPs synchronized using an ensemble meta-learning method called Rotation Forest (RF). The rejection option was retrieved for both ensembles in their model by connecting the degree of consensus from the voting majority to confidence estimations. To the best of our knowledge

3 Proposed Methodology

AVs need to drive safely through the roads. To avoid collisions, they must identify various objects around them with utmost certainty. These objects can be (other vehicles, obstacles, traffic lights, pedestrians, cyclists, driving spaces, etc.). Therefore we need a method that detects these objects with higher accuracies. Recent crashes of Tesla, Toyota, and Google self-driving cars point out that a lot needed to be done to make object detection methods for Avs. Hence there is massive scope for improvement, which must be exploited.

In this work, we proposed an object detection method for autonomous vehicles. It is a multiclass problem where the proposed method will classify between four types of objects: drivable area, other vehicles, lanes markings, and background (footpath, shops, obstacles, trees, etc.). The proposed method will consist of two phases. In Phase-I, we will classify the context. Here context can be anything (day, night, shadow, extreme light, rain, snowfall, sand storm, etc.). In the proposed method, we are taking 3 contexts which are: (1) dark conditions, (2) extreme light conditions, and (3) normal conditions. We trained a deep learning model to predict contexts using the grey-scale image to reduce computation at this phase. In Phase II, we use a model trained to detect objects specific to a particular context.

One important thing about using a context-based model is that we do not need to train deep learning models on massive datasets. Instead, we will only train the model on more minor (context-specific data). Thus our model will be lightweight because it will take less training time and loading time. So for each context, we have a separate object prediction model in the proceeding sub-sections. There are three classes we are predicting using the proposed context-based deep learning method, which: road class (R), lane marking class (L), and other object class (O), which includes footpaths, obstacles, traffic signs, traffic lights, vehicles, etc. It is an image segmentation problem.

3.1 Experimental Setup

For the validation of the proposed method, we used R programming platform. We used R packages such as for deep learning, we used the H2O package [26], for data visualization, we used ggplot2 [27] and cowplot [28] or results evaluation, we used caret package [29]. We used i7 system with 1TB hard drive, 16 GB RAM.

3.2 Data Preparation

We used the KITTI autonomous driving image dataset [30]. Every colored image is of size 1242 x 375 pixels. We create a dataset with nine attributes: r, g, b, x, y, h, s, v, e, and class. We used the Raster package [31] in R, to compute pixel values and the location of each pixel in the image frame.

Our dataset contains 13.902 million rows, so there is hardly any question of underfitting. However, we use parameter stopping metric for overfitting during the model training. We stop the model training as overfitting start. Overfitting is identified when the model stops converging, and accuracy starts to constant or decrease. So stopping the metric parameter helped us to avoid any kind of overfitting.

One of the most challenging parts was how to run on this much data our method as we have only 16 GB RAM. It was a tricky thing to manage. So, we used our RAM very cleverly. For example, we deleted unwanted variables once their use was over in the code using the rm() command. We only kept variables that are in use in RAM. Through this, we were able to manage RAM efficiently.

We compared the proposed method with the classical deep learning method. We cannot compare the proposed method with any other research paper directly because we do not know which image frames they used, and it is impossible to get the code for their method.

3.3 Proposed Method

In Figure 2. We have stated the block diagram of the proposed model. Later in the coming sections, we will use a multilayer perceptron (MLP), a feed-forward artificial neural network model that maps input data sets to appropriate outputs [26]. This is also called a fully-connected feed-forward ANN. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one.

Each node is a neuron (or processing element) with a nonlinear activation function, except for the input nodes. MLP utilizes a supervised learning technique called backpropagation for training the network. MLP modifies the standard linear perceptron and can distinguish data that are not linearly separable. Table 1 mentions some of the critical parameters configuration for the deep learning models. We need to classify three classes which include road class (R), lane marking class (L), and other object class (O).

Table 1. Some important deep learning parameters configuration for deep learning models.

No.	Parameter	Value
1.	Epochs	40
2.	Activation function	Tanh
3.	Hidden layers	2
4.	Neurons in hidden layers	C(32,32)
5.	Balance classes	True
6.	Rho	0.99
7.	momentum_start	0.0
8.	L1 regularization	0.0
9.	L2 regularization	0.0
10.	Rate	0.005

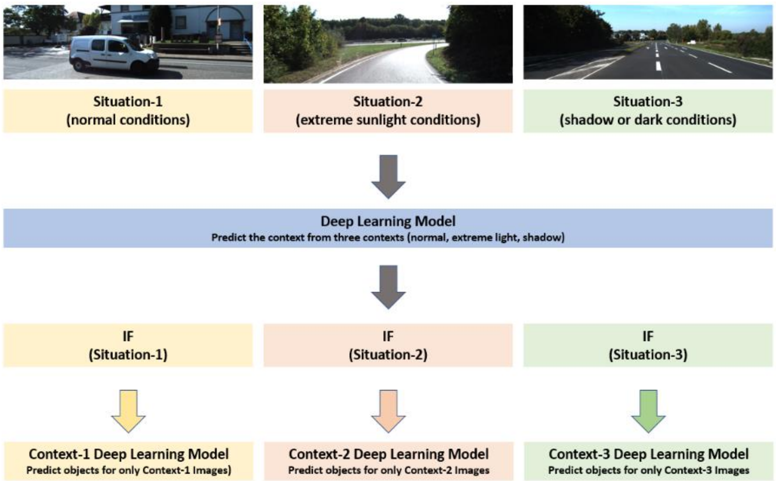


Figure 2. Proposed Method.

4 Results

For the performance evaluation of the proposed method, we used performance measurement benchmarks given by (Sokolova & Lapalme, 2009): prediction accuracy, sensitivity, and specificity. Sokolova and Lapalme analyzed 24 performance measures used in the Machine Learning classification tasks, including binary, multiclass, multi-labeled, and hierarchical. In this work, we are dealing with a multiclass prediction problem. We selected prediction accuracy, sensitivity, and specificity.

A confusion matrix is a table that shows actual versus predicted data labels. The sum of the diagonal (SoD) of the confusion matrix represents the correctly classified data label, thus can be used to compute classifier accuracy too as depicted in Figure 3, the equation which can be given as:

Accuracy% = (SoD/ Sum of all cells of confusion matrix)*100

	Actual Class A	Actual Class B
Predicted Class A	True Negative	False Positive
Predicted Class B	False Negative	True Positive

Figure 3. Sample Confusion Matrix.

We compared the classical deep learning model with the proposed context-based deep learning model. The average prediction accuracy for ten classical deep-learning image frames is 89.021%. The average prediction accuracy for ten image frames of the proposed context-based deep learning model is 95.344%. We can see that the proposed context-based deep learning model produced 6.323% better accuracy than the base scheme, as depicted in Figure 4. This proves that a context-based approach can improve object detection accuracy for AVs.

Sensitivity can be defined as the proportion of actual class labels correctly predicted by the classifier. In contrast, specificity is the ability of the classifier to identify negative results. Important terms used to calculate sensitivity and specificity are the number of true positives (TP), number of true negatives (TN), number of false-positive (FP), and number of false negatives (FN), respectively.

The proposed context-based method produced average sensitivity for ten image frames for class O of 0.9892, class L of 0.7184, and class R of 9226. The same is depicted in Figure 5. Further, in terms of specificity proposed context-based method proved average sensitivity for ten image frames for class O of 0.9113, for class L of 0.9817, and class R of 9889. The same is depicted in Figure 6.

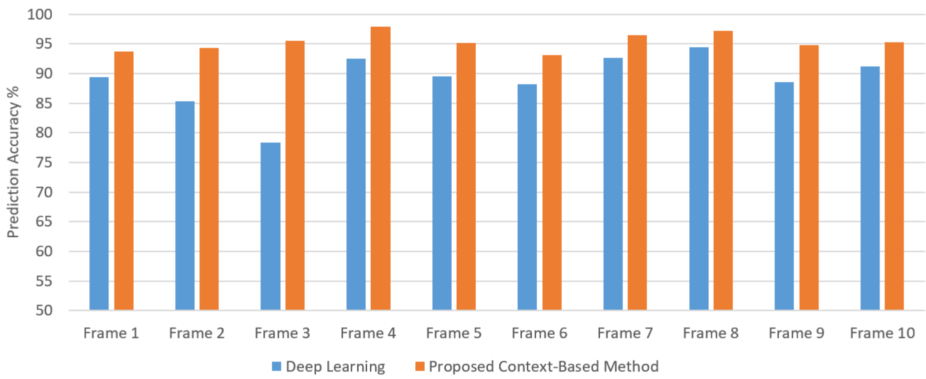
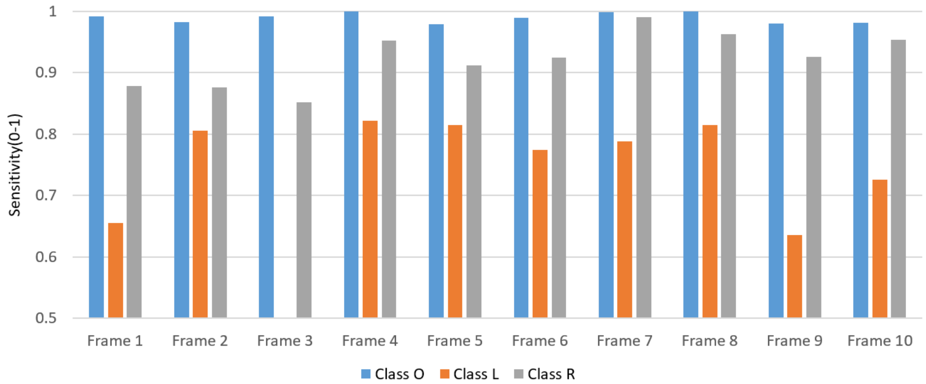


Figure 4. Prediction accuracy.

Figure 5. Sensitivity of the proposed method.



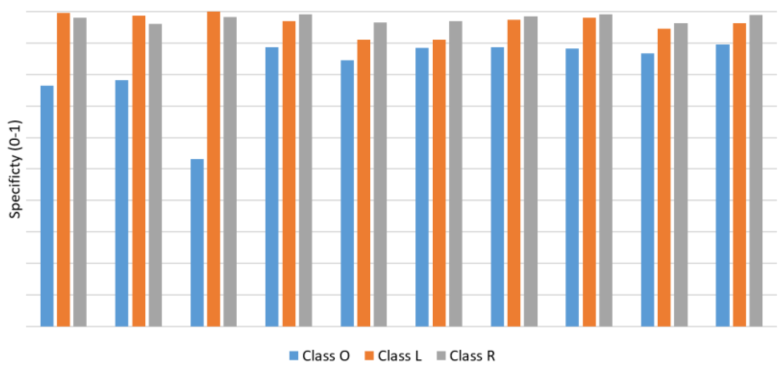


Figure 6. Specificity of the proposed method.

5 Conclusion

In this paper, we addressed the driving environment perception problem. AVs need to drive safely through the roads. To avoid collisions, they need to identify various objects around them accurately. These objects can be (other vehicles, obstacles, traffic lights, pedestrians, cyclists, driving spaces, etc.). Therefore, we need a method that detects these objects with higher accuracy. Recent crashes of Tesla, Toyota, and Google self-driving cars point out that a lot needed to be done to make object detection methods for Avs. Hence there is massive scope for improvement, which must be exploited. In this paper, we addressed object detection for AVs. The proposed object detection is a multiclass image segmentation problem. We used deep learning-based methods. The proposed method can be divided into (1) the identification of context and (2) using context-based models for object detection. For performance evaluation, we used accuracy percentage, sensitivity, and specificity. The proposed method showed promising results at par with other schemes. Our results clearly show that the proposed context-based deep learning model produced 6.323% better accuracy than the base scheme. This proves that a context-based approach can improve object detection accuracy for AVs. In the future, we want to work extensively on object detection problems to make predictions faster. Also, in this work, we have a computation power limitation that we need to overcome by using a supercomputer in our subsequent work.

References

1. Rosique, F., Navarro, P.J., Fernández, C., Padilla, A.: A Systematic Review of Perception System and Simulators for Autonomous Vehicles Research. Sensors. 19, 29 (2019). https://doi.org/10.3390/s19030648.
2. Gambino, A., Shyam Sundar, S.: Acceptance of self-driving cars: Does their posthuman ability make them more eerie or more desirable? Conference on Human Factors in Computing Systems - Proceedings. (2019). https://doi.org/10.1145/3290607.3312870.
3. Hull, D.: Tesla Is Testing Self-Driving Cars on California Roads, https://www.forbes.com/sites/alanohnsman/2017/04/25/waymo-taps-phoenix-for-a-big-public-test-of-google-self-driving-car-tech/#1ba613c06141%0A, (2017).
4. Davis, A.: Google's Self-Driving Car Caused Its First Crash, https://www.wired.com/2016/02/googles-self-driving-car-may-caused-first-crash/, (2016).
5. Shivdas, S., Kelly, T.: Toyota halts all self-driving e-Palette vehicles after Olympic village accident, https://www.reuters.com/business/autos-transportation/toyota-halts-all-self-driving-e-pallete-vehicles-after-olympic-village-accident-2021-08-27/, (2021).
6. Liane Yvkoff: The Success Of Autonomous Vehicles Hinges On Smart Cities. Inrix Is Making It Easier To Build Them, https://www.forbes.com/sites/lianeyvkoff/2020/10/28/the-success-of-autonomous-vehicles-hinges-on-smart-cities-inrix-is-making-it-easier-to-build-them/?sh=196637252bfe, last accessed 2022/03/13.

7. Arnold, E., Al-Jarrah, O.Y., Dianati, M., Fallah, S., Oxtoby, D., Mouzakitis, A.: A Survey on 3D Object Detection Methods for Autonomous Driving Applications. IEEE Transactions on Intelligent Transportation Systems. 20, 3782–3795 (2019). https://doi.org/10.1109/TITS.2019.2892405.
8. Haris, M., Glowacz, A.: Road Object Detection: A Comparative Study of Deep Learning-Based Algorithms. Electronics 2021, Vol. 10, Page 1932. 10, 1932 (2021). https://doi.org/10.3390/ELECTRONICS10161932.
9. Qian, R., Lai, X., Li, X.: 3D Object Detection for Autonomous Driving: A Survey. (2021).
10. Uçar, A., Demir, Y., Güzeliş, C.: Object recognition and detection with deep learning for autonomous driving applications: http://dx.doi.org/10.1177/0037549717709932. 93, 759–769 (2017). https://doi.org/10.1177/0037549717709932.
11. Masmoudi, M., Ghazzai, H., Frikha, M., Massoud, Y.: Object detection learning techniques for autonomous vehicle applications. 2019 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2019. (2019). https://doi.org/10.1109/ICVES.2019.8906437.
12. Alam, F., Mehmood, R., Katib, I., Altowaijri, S., Albeshri, A.: TAAWUN: A Decision Fusion and Feature Specific Road Detection Approach for Connected Autonomous Vehicles. Mobile Networks and Applications. (2019). https://doi.org/10.1007/s11036-019-01319-2.
13. Alam, F., Almaghthawi, A., Katib, I., Albeshri, A., Mehmood, R.: iResponse: An AI and IoT-Enabled Framework for Autonomous COVID-19 Pandemic Management. Sustainability (Switzerland). 13, 3797 (2021). https://doi.org/10.3390/su13073797.
14. Fujiyoshi, H., Hirakawa, T., Yamashita, T.: Deep learning-based image recognition for autonomous driving. IATSS Research. 43, 244–252 (2019). https://doi.org/10.1016/J.IATSSR.2019.11.008.
15. al Duhayyim, M., Al-Wesabi, F.N., Hilal, A.M., Hamza, M.A., Goel, S., Gupta, D., Khanna, A.: Design of robust deep learning-based object detection and classification model for autonomous driving applications. Soft Computing 2022. 1–12 (2022). https://doi.org/10.1007/S00500-021-06706-0.
16. Park, M., Kim, H., Park, S.: A Convolutional Neural Network-Based End-to-End Self-Driving Using LiDAR and Camera Fusion: Analysis Perspectives in a Real-World Environment. Electronics 2021, Vol. 10, Page 2608. 10, 2608 (2021). https://doi.org/10.3390/ELECTRONICS10212608.
17. Alam, F., Mehmood, R., Katib, I., Altowaijri, S., Albeshri, A.: TAAWUN: A Decision Fusion and Feature Specific Road Detection Approach for Connected Autonomous Vehicles. Mobile Networks and Applications. (2019). https://doi.org/10.1007/s11036-019-01319-2.
18. Ahmed, S., Huda, M.N., Rajbhandari, S., Saha, C., Elshaw, M., Kanarachos, S.: Pedestrian and Cyclist Detection and Intent Estimation for Autonomous Vehicles: A Survey. Applied Sciences 2019, Vol. 9, Page 2335. 9, 2335 (2019). https://doi.org/10.3390/APP9112335.
19. Wang, K., Zhou, W.: Pedestrian and cyclist detection based on deep neural network fast R-CNN: https://doi.org/10.1177/1729881419829651. 16, (2019). https://doi.org/10.1177/1729881419829651.

20. Annapareddy, N., Sahin, E., Abraham, S., Islam, M.M., Depiro, M., Iqbal, T.: A Robust Pedestrian and Cyclist Detection Method Using Thermal Images. 2021 IEEE Systems and Information Engineering Design Symposium, SIEDS 2021. (2021). https://doi.org/10.1109/SIEDS52267.2021.9483730.
21. Ahmed, S., Huda, M.N., Rajbhandari, S., Saha, C., Elshaw, M., Kanarachos, S.: Visual and Thermal Data for Pedestrian and Cyclist Detection. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 11650 LNAI, 223–234 (2019). https://doi.org/10.1007/978-3-030-25332-5_20.
22. Zhang, B.: Reliable classification of vehicle types based on cascade classifier ensembles. IEEE Transactions on Intelligent Transportation Systems. 14, (2013). https://doi.org/10.1109/TITS.2012.2213814.
23. Aytekin, B., Altuğ, E.: Increasing driving safety with a multiple vehicle detection and tracking system using ongoing vehicle shadow information. In: Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics (2010). https://doi.org/10.1109/ICSMC.2010.5641879.
24. Kiratiratanapruk, K., Siddhichai, S.: Vehicle detection and tracking for traffic monitoring system. In: IEEE Region 10 Annual International Conference, Proceedings/TENCON (2006). https://doi.org/10.1109/TENCON.2006.343888.
25. Zhaoxiang, Z., Yinghao, C., Kaiqi, H., Tieniu, T.: Real-time moving object classification with automatic scene division. In: Proceedings - International Conference on Image Processing, ICIP (2007). https://doi.org/10.1109/ICIP.2007.4379787.
26. Candel, A., LeDell, E., Parmar, V., Arora, A.: Deep Learning With H2O. H2O.ai Inc. (2018).
27. Wickham, H.: ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York (2016).
28. Wilke, C.O.: cowplot: Streamlined Plot Theme and Plot Annotations for' ggplot2, (2018).
29. Max Kuhn, Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt: Package "caret" Title Classification and Regression Training. (2021).
30. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. International Journal of Robotics Research. 32, 1231–1237 (2013). https://doi.org/10.1177/0278364913491297.
31. Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., Bevan, A., Shortridge, A., Hijmans, M.R.J.: Raster: Geographic Data Analysis and Modeling. CRAN. (2016).