Instance Segmentation for Autonomous Vehicle

Mohanapriya S^a, Natesan P^b, Mohana Saranya S^c, Indhumathi P^d, Mohanapriya S T P^e and Monisha R^f

Assistant Professor, Department of CSE, Kongu Engineering College, India. ^bProfessor, Department of CSE, Kongu Engineering College, India. ^cAssistant Professor, Department of CSE, Kongu Engineering College, India. ^{d.e.f}UG Students, Department of CSE, Kongu Engineering College, India.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: To detect the objects around an autonomous vehicle is the most important thing to drive it carefully without danger. This paper delivers a presentation on detecting and classifying the objects for assisting autonomous driving. In autonomous driving systems, the task of detecting the vehicle itself is one of the most important prerequisites to autonomous navigation. Deep learning, one of the computer vision tasks, performs object detection very effectively compared to earlier methods and this project is to segment the objects like vehicles, persons. Instance segmentation an extension of object detection is used to identify the objects in the image by assigning different labels for each instance of the objects belonging to the same class. Instance segmentation method is used for solving both object detection and semantic segmentation in parallel. In the existing system like Deep Neural Network (DNN) is used for segmenting the papers in the medical field and in other places. But when it comes to detect for autonomous vehicle it is not detecting clearly. To get better from this problem in the proposed system Instance segmentation is performed by using the method known as Mask RCNN. Compared to earlier detection approaches Mask RCNN shows improvement in detection accuracy and in time complexity.

Keywords: Instance Segmentation, Deep Learning, Autonomous Vehicle, Object Detection, Mask RCNN.

1. Introduction

Autonomous Vehicle

In the last few years there has been a rapid increase in research interest supporting the development of the autonomous vehicle [1], which is an automobile platform capable of sensing and reacting to its immediate environment in an attempt to navigate roadways without human intervention. Self driving cars and autonomous vehicles are hopeful solutions for increasing road safety, traffic problems and it also enhances the passenger's comfortness. The area of computer vision was continuously emerging to develop a better interaction between the human beings and machines. The key aspects of computer vision are object detection, instance segmentation and semantic segmentation.

Deep Learning

Deep learning [2-4] part of machine learning in artificial intelligence is capable of learning unsupervised data which does not contain any predefined labels. Artificial intelligence means making machines to think and act like humans. Deep Learning uses deep neural network for performing complex tasks like pattern recognition, autonomous vehicle [5, 6]. Deep neural network [7] is a neural network which contains a certain level of complexity in input layer, output layer and many hidden layer. They are capable of modelling and processing non-linear relationships where the algorithm of artificial neural networks is inspired by the human brain which learns from the huge quantity of data. Deep Learning is the technique for handling and analysing huge volumes of data by which the performance of the machine improves and also the machine identifies the hidden pattern present in the data. Since the machine is learning itself from the given data, it performs feature extraction and classification effectively without human intervention.

Moving Object Detection

The major hotspot of emerging computer vision is tracking motion vehicles and Video based object detection. The important part of the intelligent transportation system is detecting and tracking of motion vehicles. The key aspects of this include vehicle detection, image pre-processing and tracking the motion vehicle. Inter-frame difference method, background subtraction difference method and optical flow method are the major approach to detect the moving vehicle. As shown in the Figure 1 vehicles in videos and images can be recognized and one can easily identify the different types of vehicles in the given frame. In computer programs and algorithms it highly depends on the various types of data. It faces major challenges like weather conditions, lighting conditions, different type and shape of the vehicle plays a major role on making the process easier or tougher one.



Figure 1. Moving Vehicle Detection

In this frame the moving car is detected and a bounding box is created around the vehicle and there are various techniques to solve this problem. One of the methods is you can pick a pre-trained model and you can fine tune it .However there are many supervised learning approaches which requires labelled data to train the model. **Instance Segmentation**

The goal of segmentation [8, 9] is a task to simplify and to change the representation of an image into something which is easier to analyse. Instance segmentation is one of the parts of image segmentation which is used to identify every pixel, belonging to an instance of the object. It detects each distinct object. Figure 2 shows the working of Instance segmentation. Vehicle in an image is segmented as an individual object. It is a subtype of image segmentation. Instance segmentation, along with semantic segmentation, is one of two granularity levels of image segmentation. There may be a possibility of overlapping and confusion held between cars while detecting it with a bounding box. These problems can be avoided by using Instance segmentation.





Figure 2. Instance Segmentation

2. Related Work

In [1] image based object detection is performed using Deep Convolutional neural network (DNN) [10] which is supervised learning. In this paper two state art methods like 1. Single stage detection which consumes fast with less memory and 2. Region proposed detection is used. Also YOLO algorithm is used which is faster and used for single stage detectors along with DCNN which is used to improve detection and identify small objects. They have used the COCO [11] dataset model to train and test.

In recent years, unmanned aerial vehicles (UAV) [5] which has been used in platforms like military, commercial and entertainment purposes. Computer vision applications based on deep learning can work in real time. In this study, high-performance object detection algorithms like Faster R-CNN [12], YOLOv3 [13] and SSD were used to detect vehicles and the positions of vehicles driving in traffic places. The clustering based approach for finding the shortest path to reach the destination proposed in [17].

Vehicle detection and counting are increasing in the field of highway management. The detection of vehicles remains a challenge due to the different sizes of vehicles which affects the accuracy. A vision-based vehicle detection and counting system which provides the complete data for vehicle detection based on deep learning is proposed.

3. Existing System

Road vehicle detection is an essential part of many intelligent applications, such as speed and number of vehicles and driving assistance systems, automated smart parking systems, automated guidance vehicles and traffic statistics. Automatic vehicle assistance systems aimed at alert vehicle detection, and difficult monitoring of other vehicles about the roadside environment. In earlier systems Object detection for Autonomous Vehicle

using YOLOV3 Algorithm[13] was done detect the Object with Bounding box. DNN [7] was used in the existing system for object detection.

Disadvantages

- Less Accuracy
- Time Complexity
- Cannot detecting Moving vehicle correctly
- High cost

Instance Segmentation Using Mask RCNN Algorithm

The proposed approach which uses the state of art deep-learning combined with data to scan and detect, Segment and classify the objects and to estimate the position of the objects Present around the vehicle.

Mask RCNN Algorithm

The purpose of Mask R-CNN [14] (Mask Recurrent Convolutional Neural Network) is to solve the problem of object detection. R-CNN [15] which is a process by applying a mechanism called Selective Search [16] to extract regions of interest (ROI), where each ROI is represented by a rectangular boundary of an object in image. Depending on the scenario, there may be lots of two thousand ROIs. Each ROI is fed into a neural network to produce output. For each ROI's output, a collection of support-vector machines is used as a classifier to determine the type of object (if any) is contained within the ROI. Mask R-CNN is the extension of Faster RCNN [12] is a simple, flexible and general method for instance segmentation. In Mask RCNN for each instance good quality segmentation masks are generated along with efficient object detection. That is in Mask RCNN object masks are predicted in parallel with prediction of bounding box. Mask RCNN is compatible with all methods developed for instance segmentation and object detection and allows people to experiment quickly

In Figure 3 the workflow of our proposed system is given. The pre-trained input video was trained by frame to frame. Then it will read the frames in the video and be ready to proceed. Then it will apply the Instance segmentation algorithm. It will perform perspective Transformation which is transforming images between front view and top view which will correct the corners correctly into rectangular which was detected as a parallelogram. Then Thresholding will detect pixel value and blur image by assuming as binary value 0 and 1. After performing all operations it will detect the vehicle.



Figure 3. Work Flow of the proposed work

The input video frame was given. Finally it will show the output as shown in Figure 4 for vehicles and differentiate the segmentation with RGB colour format.



Figure 4. Output video frame

4. Results and Discussion

Dataset

In our proposed algorithm, pre-trained video dataset is used for evaluation. The dataset contains videos which was pre-trained frame by frame.

Accuracy

Accuracy is defined as the degree to which the result of a measurement, calculation, or specification that conforms to the correct value or a standard of the result. Table 1 shows the accuracy level of DNN and Mask RCNN algorithms.

Table 1. Accuracy level (%) of DNN and Mask RCNN

	1	
No. of image	Accuracy of	Accuracy of Mask PCNN (%)
frames in video	DININ (70)	WIDSK INCIVIN (70)
10	50	60
20	42	72
30	66	85
40	55	90

Figure 5 shows the accuracy level comparison of DNN and Mask RCNN. From the figure we can infer that Mask RCNN outperforms DNN



Figure 5. Accuracy level comparison of DNN and Mask RCNN

Time Complexity

Time complexity represents the number of operations of an algorithm that performs to complete its task. Table 2 shows the time complexity of DNN and Mask RCNN algorithms

Table 2. Tim	ne Complexity	(ms) of DNN	and Mask RCNN
	_		

No. of image	Time	Time	
frames in video	complexity of	complexity of	
	DNN (ms)	Mask RCNN (ms)	
10	50	35	
20	42	25	
30	48	20	
40	40	15	

Figure 6 shows the time complexity comparison of DNN and Mask RCNN. From the figure we can infer that Mask RCNN outperforms DNN



Figure 6. Time Complexity comparison of DNN and Mask RCNN

5. Conclusion

To detect the objects around the autonomous vehicles is essential to drive safely. Instance segmentation an extension of object detection is used to identify the objects in the image by assigning different labels for each instance of the objects belonging to the same class. Instance segmentation method is used for solving both object detection and semantic segmentation in parallel. A deep learning based instance segmentation method called Mask RCNN is proposed which performs very well in detecting objects around the autonomous vehicle. Mask RCNN shows enhanced instance segmentation accuracy and greatly reduces the time complexity when compared to previous state-of-the art methods. This makes Mask RCNN a better choice for detecting objects around autonomous vehicle.

References

- 1. Yurtsever E, Lambert J, Carballo A, Takeda K: A survey of autonomous driving: Common practices and emerging technologies. IEEE Access 2020, 8:58443-58469.
- 2. LeCun Y, Bengio Y, Hinton G: Deep learning. nature 2015, 521(7553):436-444.
- 3. Schmidhuber J: Deep learning in neural networks: An overview. Neural networks 2015, 61:85-117.
- 4. Tang Y: Deep learning using linear support vector machines. arXiv preprint arXiv:13060239 2013.
- Böyük M, Duvar R, Urhan O: Deep Learning Based Vehicle Detection with Images Taken from Unmanned Air Vehicle. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU): 2020. IEEE: 1-4.

- 6. Song H, Liang H, Li H, Dai Z, Yun X: Vision-based vehicle detection and counting system using deep learning in highway scenes. European Transport Research Review 2019, 11(1):1-16.
- 7. Soin A, Chahande M: Moving vehicle detection using deep neural network. In: 2017 International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT): 2017. IEEE: 1-5.
- 8. Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Garcia-Rodriguez J: A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing 2018, 70:41-65.
- 9. Li Y, Qi H, Dai J, Ji X, Wei Y: Fully convolutional instance-aware semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition: 2017. 2359-2367.
- 10. Khan A, Sohail A, Zahoora U, Qureshi AS: A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review 2020, 53(8):5455-5516.
- 11. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL: Microsoft coco: Common objects in context. In: European conference on computer vision: 2014. Springer: 740-755.
- 12. Ren S, He K, Girshick R, Sun J: Faster r-cnn: Towards real-time object detection with region proposal networks. arXiv preprint arXiv:150601497 2015.
- 13. Mohamed E, Shaker A, Rashed H, El-Sallab A, Hadhoud M: INSTA-YOLO: Real-Time Instance Segmentation. arXiv preprint arXiv:210206777 2021.
- 14. He K, Gkioxari G, Dollár P, Girshick R: Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision: 2017. 2961-2969.
- Girshick R, Donahue J, Darrell T, Malik J: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition: 2014. 580-587.
- 16. Van de Sande KE, Uijlings JR, Gevers T, Smeulders AW: Segmentation as selective search for object recognition. In: 2011 International Conference on Computer Vision: 2011. IEEE: 1879-1886.
- 17. Rajaram, A., & JOSEPH, S. (2014). CLUSTER BASED NEIGHBOR COVERAGE ROUTING SCHEME FOR MANET. Journal of Theoretical & Applied Information Technology, 68(3).