REAL TIME VEHICLE DETECTION AND CLASSIFICATION AT ROAD INTERSECTION USING YOLOv3

Swati N. Divatankar, Research Scholar, Department of Electronic Science, R & D Centre, Modern College of Arts, Commerce and Science, Pune, Maharashtra, India
***Prof. Dr. Magan P. Ghatule,** Principal and Head, Sinhagad College of Science, Ambegaon (Bk.), Pune, Maharashtra, India (* corresponding author)

Abstract
Intelligent traffic management systems deal with the difficult challenge of real-time vehicle detection and classification. At road intersection, with different types of vehicles, and with high density it becomes a challenging task to make accurate vehicle detection and classify it correctly with real-time speed. In the proposed research, YOLOv3 (You Only Look Once Version 3) algorithm has been used for detection and classification of the vehicles from the traffic video on a real-time basis which will further help to do traffic analysis at the road intersection. It performed better for the detection of occluded vehicles. The results of experiments show that vehicle detection and classification accuracy has been improved.

Keywords: intelligent traffic management, YOLOv3 algorithm, traffic analysis

Introduction
Real-time vehicles on the road are easily detectable and identifiable by humans. It is necessary to detect and classify the vehicles from real-time video to construct a real time vision-based system. Identifying the vehicles using computer vision techniques can achieve the same result. One of the challenging issues faced by intelligent traffic management systems is the detection and classification of vehicle from real time video [1]. The following challenges arise when trying to detect and classify the vehicles from real time video.

- Occlusion: Vehicles may be blocked by other vehicles or by man-made obstacles. Some vehicles won't be detected or tracked as a result.
- Lighting: Cameras can be impacted by light. Daylight, nightlight, and shadows may have an impact on cameras, causing some vehicles to go unobserved during vehicle tracking and detection.
- Changing weather: A change in the weather could affect visibility.
- Visual similarity: In terms of size, certain vehicles may have a similar appearance. For example, the classification of some vehicles may be incorrect because of the car and auto-rickshaw.

The YOLOv3 deep neural network technique has been employed in the proposed study to detect and classify vehicles from the traffic video in real time, contributing in traffic analysis at the road intersection.

Aim of the proposed research
The main objective of the proposed research is to detect and classify vehicles from real-time video of one of the roads of the intersection which can be further useful for traffic analysis and therefore simple vision-based tasks that can be completed by a computer or automated system will no longer require human labor.

Objectives of the proposed research
- To take a review of deep machine learning methods of vehicle detection and classification
- To develop YOLOv3 algorithm for vehicle detection and classification
- To discuss the results of real time video of moving vehicles using YOLOv3 algorithm

Vehicle detection and classification using deep machine learning
Researchers have developed a variety of algorithms for the detection and classification of vehicles, ranging from background subtraction techniques to classifier-based methods. The two methods of

vehicle detection and classification are traditional machine learning techniques and deep learning-based methods for detecting objects. The traditional machine learning techniques identify the vehicles based on their motion or shape. The appearance-based methods for detecting vehicles use color, texture, shadow features, haar-like features, etc. By combining haar characteristics with a histogram of oriented gradients, vehicle positions can be determined (HOG). Motion-based methods include the dynamics background modelling method and optical flow method [2]. Even though the vehicle detection and classification are precise, conventional machine learning techniques are not appropriate for real-time traffic scenes since they involve a human, require a lot of complicated processes to complete, and take a lot of time [3]. Convolutional neural network (CNN)-based deep learning algorithms have gained popularity recently due to advances in artificial intelligence and computer vision because they perform better than traditional machine learning techniques.

Deep machine learning can be carried out in two ways. The initial approach is based on Spatial Pyramid Pooling (SPP-net) and other is Region-based Convolutional Neural Networks (R-CNN), which are further divided into Fast R-CNN, Faster R-CNN, and R-FCN.The problem with these methods is slow speed and therefore they are not suitable for real time videos.

In order to increase the accuracy of vehicle detection and classification as well as speed, Single Shot Detector (SSD) and You Only Look Once (YOLO) have been included [4]. Despite being accurate and fast, the SSD is slower than YOLO. YOLO is the greatest option for speed [5].

The YOLO object identification approach is designed for processing in real-time. As a result, the YOLO algorithm, which is based on convolutional neural networks and can detect stationary vehicles, has been applied for the proposed research [6]. This approach provides extremely accurate and quick findings without sacrificing accuracy. The YOLOv3 can run in real time and detect all different kinds of vehicles. YOLO's outstanding speed, which can handle 45 frames per second and is three times quicker than SSD, is by far its greatest benefit [7]. YOLO uses an input image to predict vehicle bounding boxes and class probabilities. The coordinates of moving vehicles are recorded by an algorithm, which then forms a bounding box around that specific vehicle. Bounding boxes are weighted by probabilities. Depending on final weights detection is made. Depending on how many classes are used, objects are classified. Recurring detections of the same object are one of the key problems with object detection systems. By combining the YOLO algorithm with the non-max suppression technique, this issue can be solved. With this strategy, the bounding box with the highest probability is considered. So, YOLO only detects an object once.

The YOLOv1 model requires a lot of storage capacity because it comprises twenty-four layers of convolution, followed by two entirely interconnected layers. Only 49 objects can be detected, and it may miss objects that are too close. YOLOv2 has improved in both accuracy and speed. With the removal of fully connected layers and the use of anchor boxes to anticipate bounding boxes in place of connected layers, object recognition in YOLOv2 [8] is faster and more reliable than in YOLOv1. The third version of YOLO, called YOLOv3, exhibits improved speed and small object detection. In order to create the multi-label classification, YOLOv3 employs separate logistic classifiers for each class. YOLOv3 is able to detect 80 different classes of objects from one image. YOLOv3 features several enhancements over earlier YOLO algorithms, allowing it to be used for real-time vehicle identification.

Development YOLOv3 algorithm for vehicle detection, classification

The third generation of the YOLO algorithm has made some adaptive advancements over the first two, including: multi-label classification, various bounding box prediction, multiscale recognition, etc.[9]. The proposed YOLOv3 algorithm of vehicle detection, classification is as shown in Fig. 1.

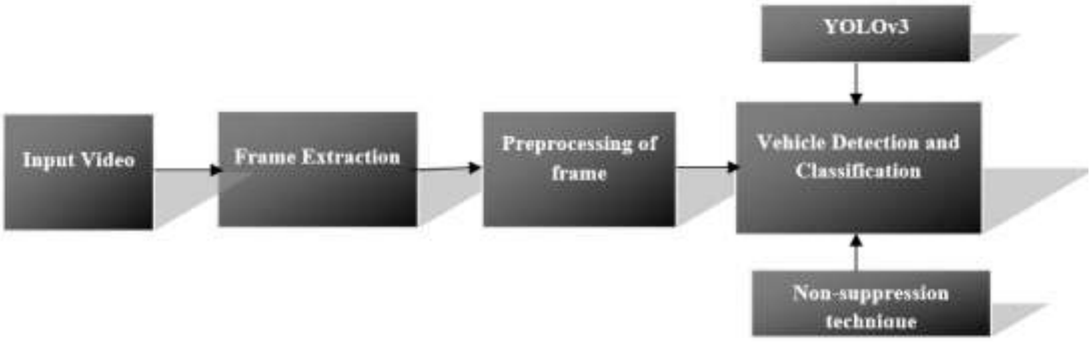

Fig. 1 Block diagram of the proposed algorithm

The system proposed is composed of three steps:

1) **Video acquisition:** Traffic video of one of the roads of the intersection is taken as an input. The video can be taken directly from the surveillance camera or recorded video. In this system, the recorded video is taken as an input. The frame is extracted from the video in RGB color space.
2) **Preprocessing of frame:** Frame is preprocessed for further process.
3) **Vehicle detection, classification using YOLOv3 and non-suppression technique:** YOLO takes input image and predicts the bounding boxes and their corresponding class probabilities for vehicles. Using non-suppression technique which is the final step of vehicle detection, the most appropriate bounding box is selected of the vehicles. The bounding box with maximum probability is considered with this technique. Multiple vehicles can be classified and localized in one frame using YOLO. YOLOv3 pretrained model is used for the research work with Open-CV. YOLOv3 is trained on COCO dataset which is trained on the 80 different classes. In this research work, five types of vehicles i.e. truck, bus, car, motorbike and bicycle are detected and classified. Therefore, from the coco dataset, the required class index holds the index of those classes.

Results and discussion:

The proposed system uses real-time recorded videos for evaluation of the performance. The real time recorded video of traffic signal of the Pune city is taken for the performance review of the proposed system. The traffic scenes consist of different vehicles car, auto-rickshaw, bus, truck, bicycle motorbike etc. The video is recorded in day time (high light intensity) with up mounted stationary CCTV camera. The recorded video is of 1280X720 frame size and 24 bit RGB color space. YOLOv3 algorithm is implemented for vehicle detection and classification using OpenCV with python language. The results for the implementations of the proposed system have been shown below in Figure 4.

Fig. 2(a)

Fig. 2(b)

Fig. 2(c)

Fig 2 Vehicle detection and classification using YOLOv3

In this approach model is trained for bicycle, car, motorbike, bus, truck etc. YOLOv3 algorithm used in the proposed approach finely detected the vehicles in the traffic scene. Some vehicles from the other side of the road are also detected and classified. To detect the vehicles from videos, first frames are extracted from video and then YOLOV3 algorithm is applied to each frame. Classified vehicles are labeled with respective labels and probability of detected class as shown in Fig 2. The algorithm shows promising results for occluded vehicles as shown in Fig. 2(a) and Fig. 2(b) as well as for moving vehicles as shown in Fig. 2(c).

Conclusion

This paper proposed YOLOv3 algorithm for vehicle detection and classification. This become challenging because of the changing environment, several vehicle models, shapes of the vehicles, and occlusion. This approach considered real time recorded traffic videos for the evaluation of the proposed system. YOLOv3 algorithm can detect and recognize the vehicles in different color boxes. It also classifies the vehicles accordingly and gives the detection probability. The proposed system shows the promising results for the occluded, near camera vehicles and for moving vehicles as well. To improve the scope of the system, the future development regarding the vehicle detection and classification requires an increase in the types of vehicle classes. Also, the vehicle images from the night vision cameras can be considered while training the model to increases the scope of the approach. The proposed research may help further for traffic analysis at the road intersection which will help intelligent traffic management systems.

References

1. Ahmad Arinaldi (2018), "Detection and classification of vehicles for traffic video analytics", INNS Conference on Big Data and Deep Learning, Procedia Computer Science 144 (2018), pp. 259–268

2. Jun Liu and Rui Zhang(2020), "Vehicle Detection and Ranging Using Two Different Focal Length Cameras", Hindawi, Journal of Sensors,Vol. 2020, Article ID 4372847

3. Jun Sang , Zhongyuan Wu, Pei Guo , Haibo Hu , Hong Xiang , Qian Zhang and Bin Cai (2018) , "An Improved YOLOv2 for Vehicle Detection", *Sensors* 2018, *18*, 4272

4. Xun Li , Yao Liu, Zhengfan Zhao, Yue Zhang, and Li He (2018), "A Deep Learning Approach of Vehicle Multitarget Detection from Traffic Video", Hindawi, Journal of Advanced Transportation,Vol. 2018, Article ID 7075814

5. Zhihao Chen, Redouane Khemmar (2019)," Real Time Object Detection, Tracking, and Distance and Motion Estimation based on Deep Learning: Application to Smart Mobility", IEEE

6. J. Redmon, S. Divvala, R. Girshick, A. Farhadi (2016), "You only look once: Unified, real-time object detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788

7. Joseph Redmon, Ali Farhadi (2018), YOLOv3: An Incremental Improvement, University of Washington

8. Joseph Redmon, Ali Farhadi(2017),"YOLO9000: Better, faster, stronger,"arXiv:1612.08242

9. Jin Liu and Dongquan Zhang (2020), "Research on Vehicle Object Detection Algorithm Based on Improved YOLOv3 Algorithm", Journal of Physics: Conference Series, 1575 (2020) 012150

