International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Detection of Multiple Objects and Analysis Using Computer Vision and Machine Learning

Susmitha Jami

Student, Rajam, Vizianagaram, 532127, India.

ABSTRACT

Numerous data are gathered by Intelligent Transportation Systems (ITS) in order to analyse the transportation system. The information can be utilized to improve the ITS, provide services for users and traffic controllers, and make transportation both safer and more effective. Computer vision (CV) techniques have been successfully applied in the visual- based ITS services due to the widespread and adaptable use of video cameras in Visual Surveillance Systems (VSS), mature edge-cloud resource scheduling for data transmission and analysis, and the quick development of deep learning. In order to better understand the relationship between the CV-based ITS services and these methods, The proposed model has focused on review the deep learning-based CV methods in the VSS and discuss the scheduling of edge-cloud surveillance resources for the CV methods in this paper. The methods include those for detection, classification, and tracking. By offering ITS services, such as traffic control, disaster management, and driver monitoring, they will address concerns and educate drivers and travelers after gathering all the features

Keywords: Computer vision (CV), deep learning, intelligent transportation system (ITS), visual surveillance system (VSS)

1. Introduction

In recent years, automatic analysis in the ITS based-on computer vision (CV) techniques has been widely applied and obtains a significant increase on its quantity with the development of computation technologies and the need of traffic services for Travellers and traffic controllers. A growing number of machine-to-machine (M2M) applications, including video surveillance, transportation, and connected vehicle applications are contributing as a majority of the growth of devices and connections, and will be 50% of the total by 2023, i.e., about 14.7 billion connections. Massive data, including visual data, can be collected from the huge urban systems by sensors and utilized for improving the performance of the services in the ITS. Intelligent transportation system (ITS) is a system which collects traffic information from vehicles and traffic facilities by sensors and analyzes them real-time to provide better services, including traffic control, disaster management, and driver monitoring for Travellers and traffic controllers in the transportation systems. By analyzing these data, we can mitigate collidance with other objects and gain information about the characteristics of that specific object.

2. Literature Survey

In paper [1]. Explains examines the edge-cloud surveillance resource scheduling for the CV methods and review the deep learning-based CV methods in the VSS, including detection, classification, and tracking methods, for a deeper comprehension of the connection between these techniques and the CV-based ITS services.

Challenges:

Identification of transportation system, visual surveillance infrastructures, data analysis center, decision center, and the services.

Proposal Methodologies:

- Numerous CV methods for visual surveillance data analysis have been proposed,
- UA-DETRAC dataset [79] is applied for evaluating the detection methods.
- Compared six widely known object detection methods, including faster R-CNN [27], SSD [29], Retina Net [60], YOLOv3[33],cascadeR-CNN [32], and FCOS [62] for vehicle detection.
- Used different methods for Vehicle Model Verification, Multi-Object Vehicle Tracking in Visual Surveillance.

In paper [2]. introduced Proposed an embedded system for traffic surveillance that can be utilized under difficulties with congestion, occlusion, and lighting night/day and day/night transitions.

Challenges:

Detection under congestion, occlusion, and lighting night/day and day/night transitions.

Proposed Methodologies:

- Specialized MF R-CNN framework to improve the detection of small-sized objects.
- Implementation of Transfer learning (VGG16) for feature extraction.

In paper [3]. proposed that there is also a method suggested for determining the animal's actual distance from the vehicle with the camera attached..

Challenges:

Animal detection on highways for preventing animal-vehicle collision using computer vision technique.

Proposed Methodologies:

- A histogram of oriented gradients (HOG) for detecting objects in a video or image.
- Cascade classifiers as it is a fast learner and requires low computation time.
- Open CV

In paper [4]. have explained Traffic scene perception (TSP) aims to extract accurate real-time on-road environment information, which involves three phases: detection of objects of interest, recognition of detected objects, and tracking of objects in motion.

Challenges:

To enhance the feature robustness to noises and image deformations.

Proposed Methodologies:

- The conventional sliding-window based method of Viola and Jones [62] practical object detector is utilized
- A clustering method is used to generate a predefined number of clusters on a specific feature space.

In paper [5]. proposes a vehicle detection technique for computer vision is proposed. The main focus of this proposed algorithm, in which the vehicle is detected based on dynamic traffic scenes. The scene can be recorded using on-board camera that fixed in position to monitor the front traffic.

Challenges:

Unwanted or false detection are gone and eliminated interference such as the road marking and surrounding objects (e.g., trees and road sign).

Proposed Methodologies:

- Video Sequence captured by computer vision
- Inside the Region of Interest, the vehicle detection procedure will be conducted (ROI)
- Blob analysis for feature recognition can more accurately and quickly identify the shadow cast by a vehicle's undercarriage.

3. Methodology

3.1 Vehicle detection using visual surveillance

In the data analysis center, Three types of CV methods are introduced into the ITS for visual surveillance data preprocessing. CV methods including detection, classification, and tracking technologies are applied for visual analysis. Then,traffic specialists in the decision center receive the analysis results from the data analysis center and real-time deal with the incidents which occur in the ITS, e.g.,traffic congestion and accidents. The UA-DETRAC dataset [79] is used to assess the detection strategies. The collection includes more than 140 000 photos that were taken from 100 movies of actual traffic in various environments. The collection contains 1.21 million bounding boxes and 8250 automobiles. The training objectives are just the vehicle bounding boxes. The effectiveness of the detection approaches is assessed using the mean average accuracy, one of the most often used assessment metrics for the detection job. For a thorough comparison, the precision vs. recall (PR) curves of several detection techniques are also included. Additionally, frame per second (FPS) and FLOPs (floating-point operations per second, in giga, G) of each detection technique are significant and contrasted in order to assess inference speed and computation quantity, which are vital elements of employment on edge devices (such as RSU in the VSS). Additionally, we provide parameter counts to demonstrate the model sizes of the approaches. Note that we built all the approaches using four Nvidia GTX1080Ti GPUs to provide for an equitable comparison. Numerous deep learning-based CV techniques for the interpretation of visual surveillance data have been presented in the last ten years. These techniques mostly concentrate on the study areas covered in Section IV. The effectiveness, the real-time, and the resource consumption of the methodologies must be assessed right away due to the growing urgency and importance of discovering accurate data analysis methods for the ITS. As an overview of these rapidly evolving research topics, we will provide three examples of the aforementioned ITS applications of vehicle detection, vehicle model verification, and multi-object vehicle tracking in this section.

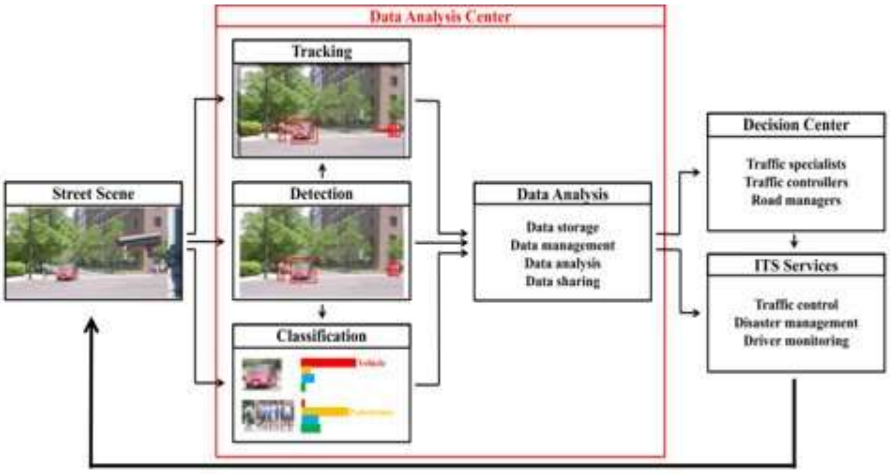
Data Analysis Center
Tracking
Street Scene
Detection
Classification
Data Analysis
Data storage
Data management
Data analysis
Data sharing
Decision Center
Traffic specialists
Traffic controllers
Road managers
ITS Services
Traffic control
Disaster management
Driver monitoring

Fig 3.1: A ITS structure

3.2 Detection methods

3.2.1: Two-stage method

The first two-stage object detection framework, Region-based Convolutional Neural Network (R-CNN) [55], generates roughly 2000 region recommendations via selective searches and feeds the ideas that have been cropped into a CNN model In order to extract features from these suggestions for classification and bounding box regression, the CNN is employed. As a result of doing a CNN forward pass for each item proposal without sharing computation, the R-CNN is sluggish. It was suggested that the R-CNN may be sped up using the spatial pyramid pooling network (SPP-Net) [25] by sharing feature maps generated from the entire set of images. Additionally, the R-restriction CNN's (fixed input image size) is removed by the spatial pyramid pooling (SPP), which also allows for the extraction of a fixed-length feature vector from the feature map for each proposal.

3.2.2 One-stage method

One-stage detection models have gained popularity in recent years because to their efficiency benefit, which more than makes up for their accuracy drawback. A fully convolutional network (FCN) was employed by OverFeat [56] to process the entire image, with each pixel of the final feature map representing a specific area in the original image. For four variants of the YOLO model, [30], [33], [35], and [58], improved modules have been included to speed up inference and increase detection precision. Similar to the YOLO concept, the single shot detector (SSD) [29] produces output maps based on a grid of anchor boxes of various scales, which are generated by many layers to detect multi-scale objects.

4. Results and Discussion

On the UA-DETRAC dataset, we compare six widely used object detection techniques—faster R x0002 CNN [27], SSD [29], RetinaNet [60], YOLOv3 [33], cascadeR-CNN [32], and FCOS [62]—to assess how well they perform on the job of visual surveillance-based object detection. Among all the detection models, the RetinaNet performs the best on mAP with an acceptable FPS (more than 50 FPS). The YOLOv3 model, however, achieves moderate mAP while obtaining the greatest FPS of 122.86 thanks to its low computational cost of 70.00 G FLOPs. In conclusion, RetinaNet can obtain the best mAP with a respectable FPS (greater than 50), making it more suitable than the others for visual surveillance object detection.

Method	Backbone	mAP (%)	FPS	FLOPs (G)	#params (M)
Faster R-CNN, 2015	ResNet101	72.7	49.02	150.66	60.13
SSD, 2016	VGG16	76.4	88.50	174.89	**24.83**
RetinaNet, 2017	ResNet101	**78.1**	53.50	143.67	55.16
YOLOv3, 2018	DarkNet53	73.8	**122.86**	**70.00**	61.54
Cascade R-CNN, 2018	ResNet101	71.5	38.40	178.45	87.93
FCOS, 2019	ResNet101	76.9	59.38	139.19	50.78

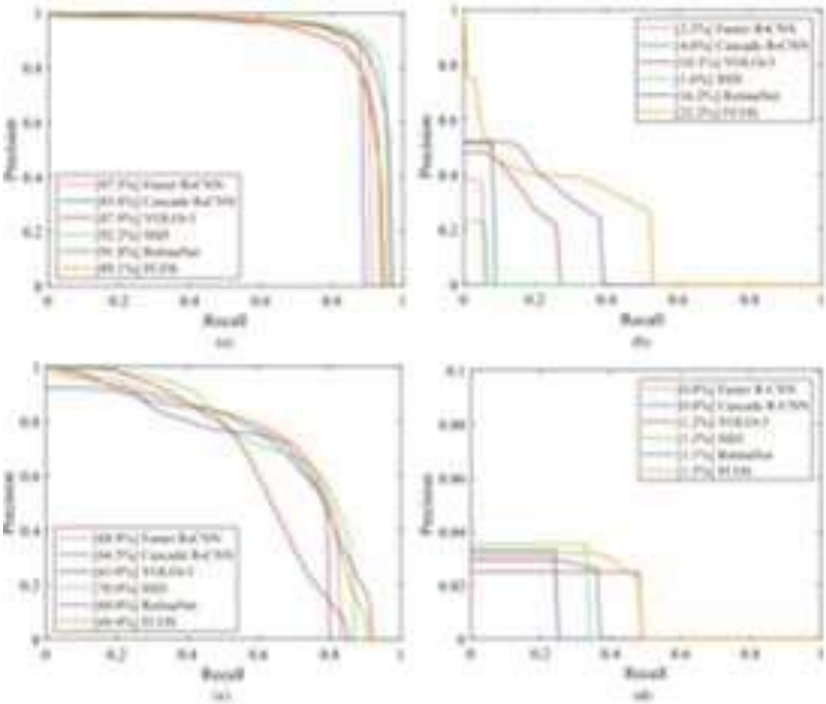

Fig:- PR curves of the detection methods in car/bus/van/others categories of the UA-DETRAC dataset. The mAP scores of the corresponding detection methodsfor the cases are reported in the legends of each subfigure to evaluate the performance of the detection methods. (a) Car. (b) Bus. (c) Van. (d) Others.

5. Conclusion

The primary goal of this essay is to improve the services offered to users of the transportation systems, such as drivers, traffic controllers, and disaster managers. We can reduce collisions with other objects and learn more about the characteristics of that particular object by analysing these data. The algorithms RetinaNet, MF customised RCNN, and HOG are more capable of detecting small objects with greater accuracy and positive rate. Every obstacle is addressed, including edge-based, nighttime, and shadow-based detection.

REFERENCES

- Xie, J., Zheng, Y., Du, R., Xiong, W., Cao, Y., Ma, Z., … Guo, J. (2021).Deep Learning- Based Computer Vision for Surveillance in ITS: Evaluation of State-of-the-Art Methods. IEEE Transactions on Vehicular Technology, 70(4), 3027–3042.
- Mhalla, A., Chateau, T., Gazzah, S., & Essoukri Ben Amara, N. (2018). An Embedded Computer-Vision System for Multi-Object Detection in Traffic Surveillance. IEEE Transactions on Intelligent Transportation Systems, 1–13.
- Sharma, S. U., & Shah, D. J. (2017). A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique. IEEE Access, 5, 347–358.
- Hu, Q., Paisitkriangkrai, S., Shen, C., van den Hengel, A., & Porikli, F. (2016). Fast Detection of Multiple Objects in Traffic Scenes With a Common Detection Framework. IEEE Transactions on Intelligent Transportation Systems.
- Nur, S. A., Ibrahim, M. M., Ali, N. M., & Nur, F. I. Y. (2016). Vehicle detection based on underneath vehicle shadow using edge features. 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)