Embedded Self Organizing Systems (Vol 10. No 7. 2023) (pp.4-8)

Embedded Selforganizing Systems

Real-time Vehicle Detection, Tracking and Counting System Based on YOLOv7

Md Abdur Rouf
Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application
Harbin University of Science and Technology
Harbin, China
abdurrouf.cse.cu@gmail.com

Qing Wu
Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application
Harbin University of Science and Technology
Harbin, China
wuqing@hrbust.edu.cn

Xiaoyu Yu
College of Electron and Information, University of Electronic Science and Technology of China, Zhongshan Institute,
Zhongshan, China
yuxy@zsc.edu.cn

Yuji Iwahori
Department of Computer Science, Chubu University,
Aichi, Japan

iwahori@isc.chubu.ac.jp

Haibin Wu*
Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application
Harbin University of Science and Technology
Harbin, China
woo@hrbust.edu.cn

Aili Wang*
Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application
Harbin University of Science and Technology
Harbin, China
aili925@hrbust.edu.cn

Abstract—**The importance of real-time vehicle detection tracking and counting system based on YOLOv7 is an important tool for monitoring traffic flow on highways. Highway traffic management, planning, and prevention rely heavily on real-time traffic monitoring technologies to avoid frequent traffic snarls, moving violations, and fatal car accidents. Three crucial duties include the detection, tracking, and counting of cars on urban roads and highways as well as the calculation of statistical traffic flow statistics (such as determining the real-time vehicles flow and how many different types of vehicles travel). Important phases in these systems include object detection, tracking, categorizing, and counting. The YOLOv7 identification method is presented to address the issues of high missed detection rates of the YOLOv7 algorithm for vehicle detection on urban highways, weak perspective perception of small targets, and insufficient feature extraction. This system aims to provide real-time monitoring of vehicles, enabling insights into traffic patterns and facilitating informed decision-making. In this paper, vehicle detecting, tracking, and counting can be calculated on real-time videos based on modified YOLOv7 with high accuracy.**

Keywords—YOLOv7, detection, tracking, counting, Google Colab, MS COCO dataset

I. Introduction

Detecting tracking and counting vehicles on urban roads and highways (such as cars, trucks, buses, motorcycles, or ambulance) and calculating statistical traffic flow information are three important tasks. The development of real-time vehicle detection, tracking, and counting systems has revolutionized traffic management and surveillance. Our research focuses on implementing an efficient and accurate system using YOLOv7, a state-of-the-art object detection algorithm. Object detection, tracking, and counting are all important steps in computer vision systems that aim to identify and quantify objects within an image or video stream. Detection involves identifying the presence of objects in an image or video frame. Tracking involves following the detected objects across successive frames in a video. Counting involves determining the number of objects of a specific class that have been detected and tracked within an image or video.

By accurately detecting tracking and counting objects computer vision systems can make informed decisions and take appropriate actions based on the information gathered from the detected objects. These techniques can be applied in various applications, such as surveillance, robotics, and autonomous vehicles, to enable these systems to operate efficiently and effectively. YOLOv7 is a popular algorithm for object detection due to several reasons such as speed and accuracy, flexibility, easy to use. However, there are other algorithms for object detection that may be more suitable for certain applications. For example, Faster R-CNN may be better suited for applications that require higher accuracy at the cost of speed, while SSD may be better suited for real-time applications that require low latency. Ultimately, the choice of algorithm for object detection depends on the specific requirements of the application, such as the desired speed, accuracy, and object categories to be detected. It is important to evaluate the performance of different algorithms on the target dataset and application requirements to make an informed choice.

A significant contribution to tools for urban/highway traffic analysis and planning [1]. A common research area in computer vision is vehicle detection and recognition, which has several potential applications in automated driving. Unfortunately, camera angles and the distance between bodies have an impact on the real-time image collection process for road vehicle onboard cameras; this results in issues including block, blur, dim light, and the tiny size of the target item. As a result, there is little acknowledgment. To increase the rate of recognition [2]. However, in this modern era of expanding technology and population, urban regions have a difficult problem with real-time highway traffic flow monitoring. Traffic rule infractions, frequent traffic jams, and fatal traffic accidents are all effects of poor road/highway traffic management. Traditional methods (such as RADAR, LiDAR, RFID, or LASAR) are time-consuming, expensive, and labor-intensive ways to solve this issue [4]. However, they fall short when it comes to classifying vehicles or gathering data on things like the number of vehicles broken down by kind and direction of travel [5].

The initial concept of such algorithm is based on the R-CNN algorithm [7]. Its fundamental design enhances the YOLO method and borrows Faster-anchor RCNN's mechanism. Similar to YOLOv3, an anchor is formed on the feature map at various scales, and a prior box is generated on the feature map for prediction [8]. The YOLO method is currently the most popular target identification technique and is used extensively in industrial manufacturing. As a result, the YOLO algorithm is the application algorithm framework employed in this paper. As stated by Wang et al. in 2022 [9], this algorithm now has the best YOLOv7 performance. Yet, there is still opportunity for advancement in the algorithm's detecting precision. On the basis of the characteristics of memory access cost, Ma et al. [10] also examined the impact of the input-output channel ratio, the number of architecture branching, and operation by element on the network reasoning speed. While scaling the model, Dollar P et al. [11] devoted more thought to activation than they did to the quantity of elements in the output tensor of the convolution layer. Some of the previous studies may lack an in-depth evaluation of the proposed systems, particularly in real-world scenarios with diverse lighting and environmental conditions. Certain studies might have focused primarily on vehicle detection and lacked comprehensive tracking and counting components. The literature review may reveal limitations in terms of handling occlusions, scale variations, or robustness to complex traffic scenarios.

The study aims to address the identified gaps and limitations in previous research by developing a real-time vehicle detection, tracking, and counting system based on YOLOv7. It proposes to incorporate novel techniques or enhancements to improve the system's performance and robustness in real-world scenarios. The study intends to provide a comprehensive solution that integrates accurate detection, robust tracking, and reliable counting of vehicles in real-time video. The proposed system has the potential to contribute to various domains, including transportation management, traffic monitoring, and surveillance applications. It can provide real-time insights into vehicle movement patterns, traffic congestion, and accurate vehicle counts, enabling better decision-making and resource allocation. Accurate detection and tracking assist in identifying potential traffic violations and enhancing overall road safety. The study's outcomes may help improve the efficiency of traffic management systems, enhance safety measures, and facilitate urban planning.

By critically assessing the literature review and identifying the gaps and limitations in previous research, your study aims to contribute to the field of real-time vehicle detection, tracking, and counting based on the YOLOv7 algorithm. The proposed enhancements, novel techniques, and comprehensive evaluation of the system can potentially address the identified limitations and provide a more accurate and robust solution for real-world scenarios.

II. Vehicle detection tracking and counting

A. Vehicle detection

Traditional machine learning algorithms for vehicle detection completely rely on hand-crafted and manually retrieved object features [12]. The look or motion characteristics of pixels in a picture are used to obtain the feature vector [13]. The recognition algorithms, which identify cars based on their visual characteristics and shape, essentially take a three-step approach. By erasing the background, they first do background modeling or background subtraction. The remaining blobs and their locations are found in the second stage, which follows the removal of the background. In the third stage, classification techniques are used to determine the different types of vehicles using the observed blobs and extracted data. For offline videos with a consistent background and favorable weather, these algorithms work effectively. However, particularly vulnerable to rapid background changes, unique weather, and issues with color and shape [14].

However, it is challenging to manually build a comprehensive feature descriptor that can accurately characterize all varieties of vehicles due to the range of partial or full occlusions, slowness, appearances, illumination, and backdrop changes. Robust and precise vehicle recognition solutions have been created as a result of the recently discovered deep learning field, resolving the problems that machine learning-based detection algorithms had in the past. Artificial neural networks (ANN), deep neural networks (DNN), and convolutional neural networks are the three main models used in these algorithms (CNN). These techniques use automatically retrieved deep and hidden properties to find automobiles. Many research projects based on the CNN architecture have been carried out. For instance, the most effective and extensively used techniques are region-based (RCNN) [15], faster region-based (FRCNN) [16], SSD (Single Shot Detector) [17], ResNet, RetinaNet, and YOLO (v1, v2, v3).

Although RCNN, FRCNN, SSD, ResNet, and RetinaNet have somewhat higher detection and classification accuracy than YOLO, only YOLO can be utilized in real-time video analysis tools like traffic monitoring due to its fast detection. While YOLO can achieve a minimum of 20 FPS and a maximum of 150 FPS, other approaches have a maximum frame rate of 8 [19]. However, YOLO accomplishes the object localization and classification tasks on an input image with a single neural network just once, hence the algorithm's car classification accuracy is less than 57%. As a result, the categorization accuracy is so low. On urban roads and highways, 57% categorization accuracy is insufficient for traffic flow extraction methods. Hence, by incorporating a reliable classifier layer within the algorithm, YOLO-based vehicle detection algorithms' classification accuracy should be increased.

B. Tracking

Traffic flow information such as vehicle speed, motion direction, total and categorical numbers, duration between entry and exit points of vehicles, and other desired parameters are highly valuable when the trajectories of vehicles (X and Y trajectory vectors) are extracted from a roadway video scene for the specific time interval or in real-time. These traffic data make it simple to keep an eye on and control the flow of traffic in any desired section of a city or urban area. Although though there are a number of vehicle tracking techniques that make use of pixel and bounding-box features, including optical flow, particle filter, and Kalman filter, none of these methods are adequate for accurately solving the vehicle tracking problem in real-time on highways [18]. Tracking algorithms, such as Kalman filtering or deep association methods, can be applied to track objects across consecutive frames based on their detected bounding boxes.

C. Counting

Object counting involves determining the number of objects of a specific class that have been detected and tracked within an image or video. This can be achieved by using object detection, tracking, and classification to count the number of objects that meet specific criteria. For example, an object counting system in a traffic video could use object detection, tracking, and classification to count the number of cars on a road. Overall, object detection, tracking, classifying, and counting are important steps in computer vision systems that can be used in a variety of applications, such as surveillance, robotics, and autonomous vehicles. Object counting can be performed by tracking the number of unique objects detected or by analyzing the flow of objects across predefined boundaries. The detection, tracking, and accurate counting will be expeditious on highways are speedy vehicles. By accurately detecting, tracking, classifying, and counting objects, these systems can make informed decisions and take appropriate actions based on the information gathered from the detected objects.

III. EXPERIMENTS AND RESULTS ANALYSIS

We carry out experiments and validate our object detection approach using the Google colaboratory on our modified MS COCO dataset. That dataset has five classes as ['Ambulance', 'Bus', 'Car', 'Motorcycle', 'Truck'] using this dataset we pre-trained the 627 images the annotations were 1194, average image size of 0.72mp and the median ratio of 1024x751. Implementing counting techniques to accurately estimate vehicle numbers in specific regions. As shown in Fig. 1, the metrics of training improvement results on TensorBoard are increasing and loss functions are decreasing.

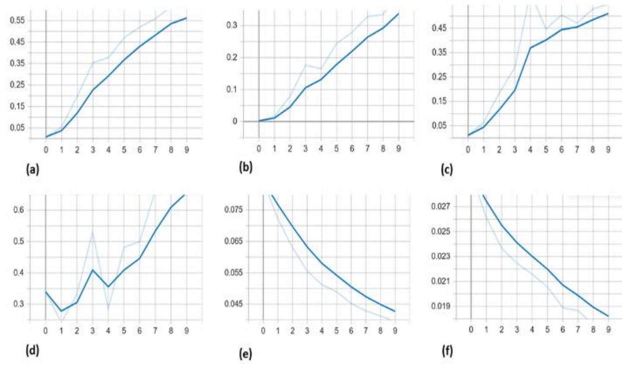

Fig. 1. Improvement training results on TensorBoard (a) metrics_mAP_0.5; (b) metrics mAP_0.5_0.95; (c) metrics_precision; (d) metrics_recall; (e) train_box_loss; (f) train_cls_loss.

Also, we have developed the dectect.py model that can be detected, tracked, counting images objects, or real-time videos. Fig. 2 shows our training images test results and Fig. 3.shows the Original model training data results. As shown in Fig. 4. Our developed model results also show Fig. 5 trained image curves. Real-time showing on the screen each FPS how many buses, cars, motorcycles, etc. Finally, we demonstrated the object detection best performance compare to the original model results on our developed model.

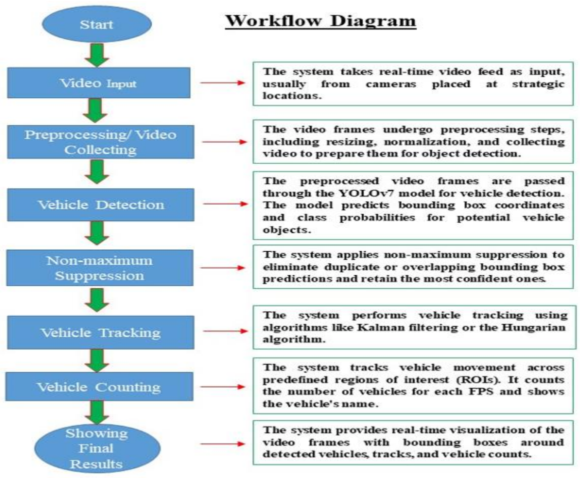

Fig. 2. Workflow diagram of system.

Fig. 3. Training images test results.

Fig. 4. Original model training data results.

Fig. 5. Developed model training data results.

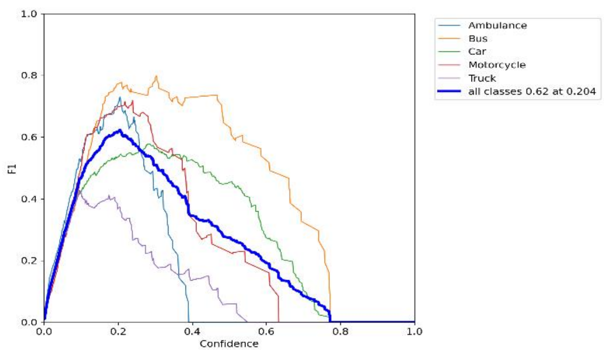

Fig. 6. New train data improved F1_curve.

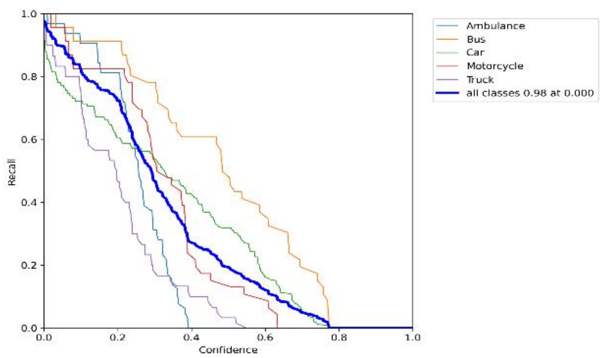

Fig. 7. New train data recall curve.

IV. CONCLUSIONS

In this paper, we proposed a real-time vehicle detection, tracking, and classifying system based on the YOLOv7 algorithm. The proposed system can accurately detect, track, and classify different types of vehicles in real-time, and can be used for monitoring traffic flow on highways. Also we are developed the dectect.py model that can be detect, tracked, counting the images objects or real time videos. The system can provide valuable insights into the traffic flow on highways, which can help transportation authorities make better decisions to improve the efficiency and safety of the transportation system. Future work includes improving the accuracy of the system by incorporating additional features such as speed estimation and lane detection.

In summary, the research on "Real-time Vehicle Detection, Tracking and Counting System Based on YOLOv7" aims to contribute to the field by addressing the limitations of previous studies and developing an enhanced solution for accurate and efficient vehicle analysis. The proposed system, leveraging YOLOv7 and incorporating novel techniques, has the potential to provide real-time insights and impact various domains, including transportation management, traffic monitoring, and surveillance applications. By improving the accuracy and robustness of vehicle detection, tracking, and counting, this research can make significant strides in advancing the field of real-time vehicle analysis.

ACKNOWLEDGMENT

Qing Wu acknowledges the National Natural Science Foundation of China (Grant No. 62205091), the China Postdoctoral Science Foundation Funded Project (Grant No. 2022M710983) and HeiLongJiang Postdoctoral Foundation (Grant No. LBH-Z22201). Supported by the Fundamental Research Foundation for Universities of Heilongjiang Province: 2022-KYYWF-0121.

REFERENCES

[1] F. Li, C.-H. Lee, C.-H. Chen, L.P. Khoo, Hybrid data-driven vigilance model in traffic control center using eye-tracking data and context data, Adv. Eng. Inform. 42 (2019) 100940, https://doi.org/10.1016/j.aei.2019.100940

[2] Zha, H.; Miao, Y.; Wang, T.; Li, Y.; Zhang, J.; Sun, W.; Feng, Z.; Kusnierek, K. Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction with Machine Learning. Remote. Sens. 2020, 12, 215.

[3] Y. Zhang, Y. Sun, Z. Wang, and Y. Jiang, "YOLOv7-RAR for Urban Vehicle Detection," Sensors, vol. 23, no. 4, Art. no. 4, Jan. 2023, doi: 10.3390/s23041801.

[4] S. Khan, H. Ali, Z. Ullah, M.F. Bulbul, An intelligent monitoring system of vehicles on highway traffic, in: 2018 12th International Conference on Open Source Systems and Technologies (ICOSST), 2018, pp. 71-75, https://doi.org/10.1109/ ICOSST.2018.8632192.

[5] Z. Zhao, P. Zheng, S. Xu, X. Wu, Object detection with deep learning: A review, IEEE Trans. Neural Netw. Learn. Syst. 30 (11) (2019) 3212-3232.

[6] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M., 2022. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696

[7] Wang, H.; Li, Z.; Ji, X.; Wang, Y.J.a.p.a. Face r-cnn. arXiv 2017, arXiv:1706.01061.

[8] Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv 2022, arXiv:2207.02696.

[9] Ma, N.; Zhang, X.; Zheng, H.-T.; Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116-131.

[10] Dollár, P.; Singh, M.; Girshick, R. Fast and accurate model scaling. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 924-932

[11] Wang, S.-H.; Fernandes, S.L.; Zhu, Z.; Zhang, Y.-D. AVNC: Attention-Based VGG-Style Network for COVID-19 Diagnosis by CBAM. IEEE Sens. J. 2021, 22, 17431-17438.

[12] H. Luo, J. Liu, W. Fang, P.E. Love, Q. Yu, Z. Lu, Real-time smart video surveillance to manage safety: A case study of a transport mega-project, Adv. Eng. Inform. 45 (2020) 101100, https://doi.org/10.1016/j.aei.2020.101100.

[13] Sivaraman, M.M. Trivedi, Looking at vehicles on the road: A survey of visionbased vehicle detection, tracking, and behavior analysis, IEEE Trans. Intell. Transp. Syst. 14 (4) (2013) 1773-1795.

[14] N.A. Mandellos, I. Keramitsoglou, C.T. Kiranoudis, A background subtraction algorithm for detecting and tracking vehicles, Expert Syst. Appl. 38 (3) (2011) 1619-1631, https://doi.org/10.1016/j.eswa.2010.07.083.

[15] R. Girshick, Fast r-cnn, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440-1448.

[16] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (6) (2017) 1137-1149.

[17] W. Liu et al., Ssd: Single shot multibox detector, in: Computer Vision -ECCV 2016, Springer International Publishing, Cham, 2016, pp. 21-37

[18] M.B. Khalkhali, A. Vahedian, H.S. Yazdi, Vehicle tracking with kalman filter using online situation assessment, Robot. Autonomous Syst. (2020) 103596, https://doi. org/10.1016/j.robot.2020.103596.