VEHICLE COUNTING AND CLASSIFICATION USING COMPUTER VISION

[1]Gaurav. T, [2] Ms. Komala R

[1]Student, [2]Professor
Master Of Computer Applications
MS RAMAIAH INSTITUTE OF TECHNOLOGY

Abstract-

Vehicle counting and classification using the YOLO (You Only Look Once) algorithm is a technique for detecting and recognizing vehicles in images or video streams in real-time. To detect objects, the YOLO algorithm breaks down the picture into grids. The YOLO algorithm has been used in various applications, including traffic surveillance and management systems, where it can accurately count and classify vehicles in real-time. The YOLO algorithm has several advantages over other object detection algorithms, including its high detection accuracy, fast processing speed, and the ability to handle multiple objects in a single image. However, the YOLO algorithm also has some limitations, such as the difficulty in detecting small objects and objects that are partially occluded. Overall, vehicle counting and classification using the YOLO algorithm is a promising field of research that has the potential to enhance traffic management and improve safety on roads.
While the algorithm has some limitations, its high detection accuracy and fast processing speed make it a suitable option for real-time applications. Further development has the potential to lead to improvements in this field which would help in better traffic control and management.

Keywords: YOLO algorithm, real-time processing, traffic surveillance, neural network, object detection

INTRODUCTION

Vision-based vehicle counting is a technology that uses image processing and computer vision techniques to count the number of vehicles passing through a particular area. It is a widely used method for traffic monitoring and management in urban areas, highways, and parking lots.

Vehicle counting can be considered as process to estimate the density of traffic and assess them for intelligent transportation systems. With a lot of improvement in technology, cameras are installed in most of the roads.

Deep learning has revolutionized the field of vehicle detection by providing accurate and efficient methods for identifying and locating vehicles in images and videos. One of the advantages of deep learning-based vehicle detection is its ability to learn features directly from the data. This makes deep learning-based methods more efficient as it can handle changes in various viewpoint and lightings. Another advantage of deep learning-based vehicle detection is its ability to handle real-time applications.

The system captures images or video footage of the passing vehicles and then analyzes the data to identify and count the vehicles. This technology can be used to measure traffic flow, detect congestion, and provide useful insights to traffic engineers and planners for optimizing traffic management strategies. With the advancement in technology, vision-based vehicle counting systems are becoming increasingly accurate and reliable, providing real-time data for efficient traffic management.

Because everything can be tracked from a specific location, such as how many vehicles have crossed this toll and whose vehicle it was, the traffic police might benefit from a vehicle detection and counting system. Individuals find it challenging to film every car in their vicinity because they are driving by in real-time. This application is extremely capable of obtaining the time-saving quality and being automated in order to get over this limitation. Given that this application only requires a camera and a few wires and can be deployed anywhere to establish connectivity with the central system. An officer can monitor heavy traffic in that area and transmit updates to the following toll officer so that they are prepared for it whenever it occurs.

LITERATURE SURVEY

Object detection has been the subject of extensive study. Ross Girshick, Jian Sun, Shaoqing Ren, and Kaiming He advocate using the Regional Proposal Network in their paper [1]. In order to practically enable cost-free region proposals, they devised a Region Proposal

Network (RPN) that shares full-image convolutional features with the detection network. Fully convolutional networks, also known as RPNs, forecast object limits and objectness scores for each point simultaneously. In an effort to boost performance, they combined the convolutional elements of RPN and Fast R-CNN into a single network. The training and testing phases both made use of the PASCAL VOC dataset. The Microsoft COCO object identification dataset's 80 different item categories have also been used by them.

In the paper titled Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Identification [2], the authors looked at three different object detection methods to determine the type of vehicle. Although the Faster-RCNN model is the fastest among the RCNN models, it fails to provide an acceptable FPS as it uses CNN. The SSD (Single-shot detector) is rapid but uses mobile-v1, which decreases accuracy. However, in YOLO version 4, the FPN was used to forecast features for every layer of the YOLO framework (Feature Pyramid Network). The recognition of high-resolution properties by YOLO resolved the issue of not catching little objects. With just one GPU, Version 4 was developed to improve object detection.

We can see from [3] that they were successful in trying to make the YOLOv3 algorithm better. The original Yolo v3 algorithm estimates the initial width and height of the expected bounding boxes using the k-means cluster method. Since the projected breadth and height depend on the original cluster centers, this method takes a long time to handle huge datasets.
The unique method offers a faster convergence time for selecting more representative initial widths and heights for the anticipated bounding boxes, according to their simulation results. For the MS COCO dataset, their suggested technique had a higher average IOU and performed more quickly. The typical IOU is 60.44 percent, which is 0.5 percent greater than the first figure made using the YOLOv3 technique. Moreover, only 1/199 of the operational time of the original YOLOv3 technique is now in use. The average IOU is 67.45%, which is 0.13% more than the original YOLOv3 method, and the PASCAL VOC dataset runs in 1/81 the time of the original YOLOv3 method. In terms of recall, average mean precision, and F1-score, their proposed solution outperforms the original YOLOv3 method.

In another paper [4], they presented a framework for vehicle detection using convolution neural network. They used feature fusion approaches to combine high-level and low-level features to detect various vehicle sizes on various attributes. They indicated that because of changes in lighting, background clutter, occlusion, etc., conventional vehicle recognition techniques like the Gaussian Mixed Model (GMM) were not ideal. To group the dataset and discover prior knowledge, they employed k-means. Moreover, feature concatenation was used to extract more detailed features. Their architecture significantly outperforms Faster R-CNN and SSD, especially for tiny automobiles, with the help of these technological advancements. The speed of their network was found to be 15 FPS, which is three times faster than the Faster R-CNN.

In the approach they offered in [5], vertices and their topological structure are seen as the basic characteristics for classifying vehicles. To identify autos, they used a classifier based on multi-layer perceptron networks (MLPN). The application of methods depend on the gradient descent approach with lowest exponential function error is utilized in this neural network classifier as learning. According to experimental findings, a parameterized model using a neural networks classifier can sufficiently and accurately describe automobiles with an accuracy rate of greater than 91%. Sadly, there is a flaw in their suggested model. The test was conducted in a toll booth. We don't yet know how the model would fit in other traffic scenarios because they have not yet tested how well it works in those cases.

In [7], they used handheld cameras to capture videos on a highway road. YOLO algorithm was used for object detection purpose which attained a remarkable outcome. Using the YOLO framework's generated bounding boxes, they constructed multiple object tracking using correlation filters. The proposed method can accurately recognize, track, and count the cars, according to experimental analysis of real video sequences. Although the model provided efficient results when tested, the processing time it took for object detection was significantly long.

The authors of [9] used the Active Basis Model (ABM) to categorize automobiles. Sequential frame vehicle detection is provided to the counting process, and counted vehicle classification is provided to the counted vehicle. To train an RF and divide vehicles into three categories, such as small (a vehicle like a car), medium (such as a van), and large, the automobile length in the TSI image and the correlation value produced from the GLCM matrix are employed (like a bus and trunk).

METHODOLOGY

You Only Look Once is a shorthand for YOLO. It is an algorithm for real-time object recognition. Several objects can be classified and localized in one frame.

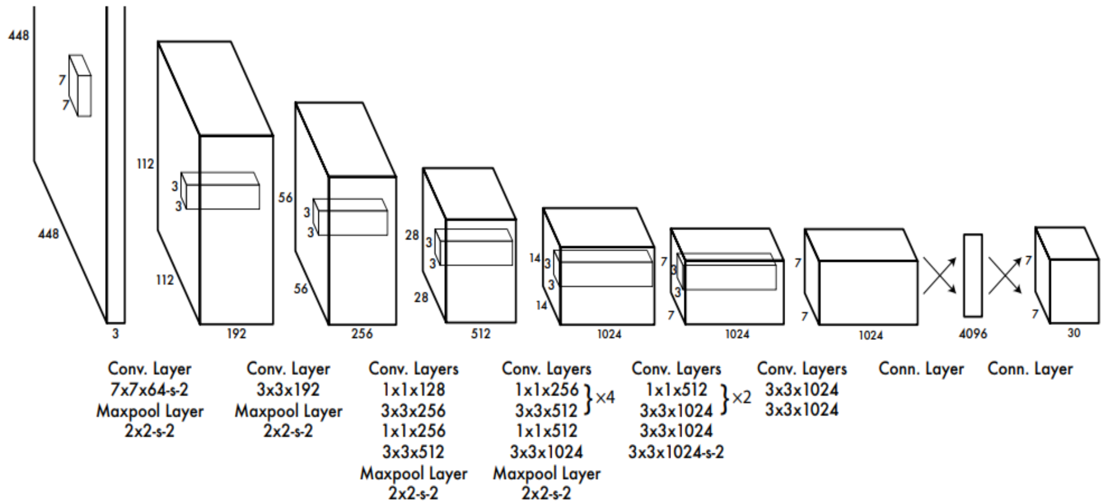

Fig. 1. The architecture of Residual Neural Network model from the paper [11]

YOLO Architecture explanation:

There are 24 convolutional layers in the YOLO network, followed by 2 fully linked layers. Pre-trained on the ImageNet classification task at half resolution (224 224 input picture), the convolutional layers are subsequently trained on detection at double resolution.
To condense the feature space from earlier layers, use layers that alternate between 11 reduction layer and 33 convolutional layers.
To train the network for object detection, the final four layers are added.
The last layer forecasts the likelihood of an item class and a bounding box

Benefits of YOLO:

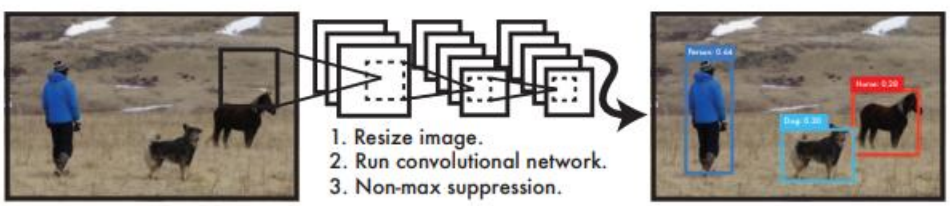

Fig. 2. The YOLO detection system from the paper [11]

- YOLO is simple: Figure 2 shows how wonderfully easy YOLO is. Several bounding boxes and class probabilities for each box are predicted by a single neural network. The detecting performance of YOLO is immediately enhanced while training on whole images. When compared to conventional item identification techniques, this unified approach offers a number of advantages [1].
- YOLO is quite quick: Since detection is framed as a regression problem, a complicated pipeline is not necessary.
- YOLO discovers universal representations of items: When tested on creative shots and trained on real-world images, YOLO greatly outperforms leading detection methods like DPM and R-CNN. It is less likely to fail when employed in unfamiliar

circumstances or with unexpected inputs because YOLO is so generalizable.

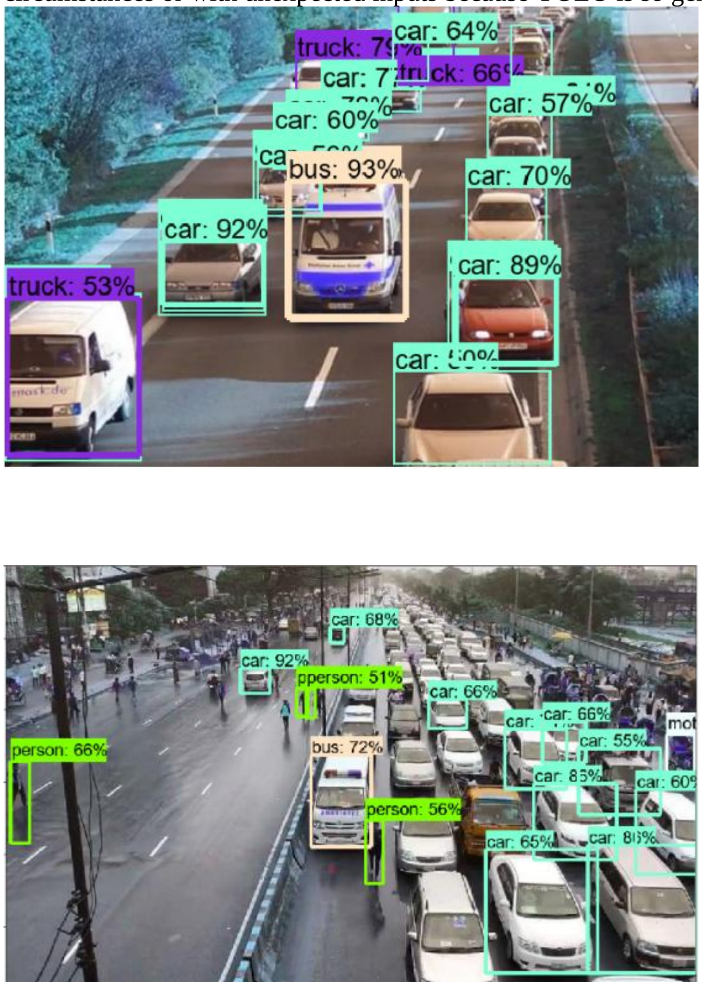

Fig. 3. Examples of car detection by YOLO from the paper [12]

Vehicle Classification:

There are three sizes of vehicles: small (like a car), medium (like a van), and enormous (for example, bus and trunk). To do this, we take two characteristics out of each vehicle type that let us discriminate between them. Initially, we develop a length-based feature that is very effective in distinguishing between various car sizes.

Training dataset: Training can be carried out using a variety of video clips shot on highways at various times.

CONCLUSION

In conclusion, the YOLO algorithm has shown promising results in vehicle counting and classification in various traffic scenarios. The real-time object detection capability of YOLO and its ability to generate bounding boxes make it a suitable choice for detecting and tracking multiple vehicles in a scene. However, the processing time for object detection can be a limitation, especially in scenarios with heavy traffic. Researchers have proposed different techniques to address this issue, for example, categorizing vehicles according to size and length or applying correlation filters for object tracking. Overall, the use of YOLO for vehicle counting and classification has proven to be an effective approach, but further research is needed to improve its accuracy and efficiency in various traffic situations.

REFERENCES

[1] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031

[2] J. -a. Kim, J. -Y. Sung and S. -h. Park, "Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition," 2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia), Seoul, Korea (South), 2020, pp. 1-4, doi: 10.1109/ICCE-Asia49877.2020.9277040

[3] Zhao L, Li S. Object Detection Algorithm Based on Improved YOLOv3. Electronics. 2020; 9(3):537. https://doi.org/10.3390/electronics9030537

[4] Chen, L., Ye, F., Ruan, Y. et al. An algorithm for highway vehicle detection based on convolutional neural network. J Image Video Proc. 2018, 109 (2018). https://doi.org/10.1186/s13640-018-0350-2

[5] Wei Wu, Zhang QiSen and Wang Mingjun, "A method of vehicle classification using models and neural networks," IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No.01CH37202), Rhodes, Greece, 2001, pp. 3022-3026 vol.4, doi: 10.1109/VETECS.2001.944158.

[6] W. Balid, H. Tafish and H. H. Refai, "Intelligent Vehicle Counting and Classification Sensor for Real-Time Traffic Surveillance," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 6, pp. 1784-1794, June 2018, doi: 10.1109/TITS.2017.2741507.

[7] C. S. Asha and A. V. Narasimhadhan, "Vehicle Counting for Traffic Management System using YOLO and Correlation Filter," 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2018, pp. 1-6, doi: 10.1109/CONECCT.2018.8482380.

[8] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui and P. Sanchez, "Adaptive Multicue Background Subtraction for Robust Vehicle Counting and Classification," in IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 2, pp. 527-540, June 2012, doi: 10.1109/TITS.2011.2174358.

[9] Kamkar, Shiva & Safabakhsh, Reza. (2016). Vehicle detection, counting and classification in various conditions. IET Intelligent Transport Systems. 10. 10.1049/iet-its.2015.0157.

[10] Song, H., Liang, H., Li, H. et al. Vision-based vehicle detection and counting system using deep learning in highway scenes. Eur. Transp. Res. Rev. 11, 51 (2019). https://doi.org/10.1186/s12544-019-0390-4

[11] Redmon, Joseph & Divvala, Santosh & Girshick, Ross & Farhadi, Ali. (2016). You Only Look Once: Unified, Real-Time Object Detection. 779-788. 10.1109/CVPR.2016.91.

[12] Roy, Shuvendu. (2019). Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network. 10.1109/ECACE.2019.8679295.

[13] Shaoyong Yu, Yun Wu, Wei Li, Zhijun Song, Wenhua Zeng, A model for fine-grained vehicle classification based on deep learning, Neurocomputing, Volume 257, 2017, Pages 97-103, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2016.09.116.