Vehicle Detection,Classification & Counting using YOLOv3

Sripathi Pavan Kumar, Korra Srinu, Yearva Sai Kushal Reddy, Saroja kumar Rout and Farhana Begum

Department Of Information Technology, Vardhaman College of Engineering, Shamshabad, Hyderabad, 501218, Telangana, India.

*Corresponding author(s). E-mail(s): pavansripathi21@gmail.com; Contributing authors: korrasrinuchauhan@gmail.com; kushalreddy1010@gmail.com; rout_sarojkumar@vardhaman.org; farhanasattar@vardhaman.org;

Abstract

Traffic tracking is one vicinity that makes use of deep getting to know for some of purposes. By the usage of cameras hooked up at positive places at the roads, many duties including car counting, car identification, visitors violation tracking, car pace tracking, etc.. can be solved. The first step for automobile disclosure in YOLO v3 is to train a cnn with a large image collection containing vehicles can be done through superintended research where the CNN is fed a set of labeled images and learns to notice the presence and location of vehicles in the images.Once qualified CNN can be utilized to detect automobiles in real time by inputting images or video frames and outputting the detected vehicle position and class probabilities to the video. The orchestrates of the moving objects gravitational pull on the count line in different sections of the road its amount of motorcycles small cars and large vehicles is recorded the actual number of vehicles on the road is compared with the number of vehicles measured by the system using a confusion matrix three different counting periods were used to figure out the result.

Keywords: Convolutional Neural Network (CNN), Region-Based Convolutional Neural Network (RCNN), Deep Learning (DL), DarkNet-53, You Only Look Once - version 3 (YOLOv3).Common Objects in Context (COCO)

1 Introduction

Vehicle realisation is an important aspect of cv with various fields such as driverless cars traffic monitoring and collision prevention systems it is the detection and location of automobiles in videos or photographs convolutional neural network one common approach for vehicle detection is to use CNN which are the default parameters of ann designed to process data with a matrix geometry such as images yolov3 is a popular cnn for vehicle detection that you only need to look at once.

This data can be utilized in traffic management to enhance traffic flow. Sensor technology has progressed dramatically in recent years. With progressions in micro-controllers various sensors have become affordable enough to be utilized in photodiodes diminishing costs and time the vehicle detector process is divided into three parts. The detector is the first component. Acoustic sensors, magnetometers, accelerometers, ultrasonic and microwave radars, and optical scans are examples of sensors. Signal processing is the second component.[5] For instance, consider the discrete wavelet serial correlation functional method. The third element is data preparation and analysis, such like Location information for moveable cars. The most frequently employed method of evaluating throughput is premised on restrictive detection that either encompasses scanner beneath the asphalt such a system necessitates regular maintenance and validation exacerbating traffic chaos.

Outdoor CCTV image recognition seems to be more delicate to climate brightness and gloominess than other systems object recognition devices in contrast side might provide significant advantages which include not trying to disrupt downtown areas ease of installation and ease of alteration. In the past century vehicle sensor research has grown significantly enabling for the explosive growth of smart transportation systems the progression of technology for image processing and also the widespread deployment of road cams have crafted image-based vehicle recognition and classification smoother to discover simple motion information its most popular media image identification system is to use feature extraction.[6]

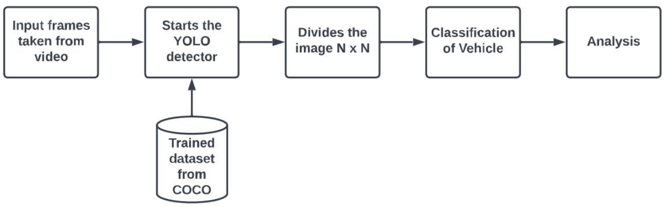

Fig. 1 Working Model of VDCC

To extract the picture of the vehicle on the street, the quantitative change between the developed baseline sequence as well as the current video frame is assessed for each incoming traffic bounding box. To determine the vehicle's foreground zones, a supervised learning technique, such as a Gaussian framework structure, must be trained and preserved.For the object to be detected, each of the preceding approaches require a stable background.[7] The application has a tough time handling balling and blockage of huge cars, often causes numerous vehicles to appear as one entity. A non-static adaptive background is required to instant activity detecting. This automated facilitates in monitoring and forecasting organic traffic inside any region. As a result, it is one of the critical approaches for optimising traffic signals. There seem to be numerous ways for estimating vehicle load, including interferometry or initiation loops, radars, and traffic cameras.

The following is the way the paper is structured: The Literature Review is described in Section 2. Section 3 introduces the envisioned vehicle detection and classification system. Section 4 summarizes the findings and discussion. Section 5 concludes with conclusions and further work.

2 Literature Survey

There have been many studies on vehicle detection using YOLOv3 and other CNN-based approaches in the literature. The data originates from two different varieties of sensors: active and passive sensors. The video feed is generated on pictures from cameras as well as satellites. RCNN models have some drawbacks, which can be solved by using a fusion of YOLO and rapid R-CNN. Here they have done detection and classification using CNN and have shown the results on screen only, with accuracy of 70% .[1]

In this work, Gabor wavelets and local adaptations The multi - resolution vehicle features are extracted using a classification algorithm operator.Then the image is split into several smaller ones to extract the weighted region The histogram sequence is ready and focused on the correct image View vehicle features. Main ingredient Employing analytics to achieve low dimensionality Histogram function that measures the closest reading Probabilities from neighbor weights in the traffic data set for the final score.Experimental evidence is Over 97%, only 3% wrong rate, 91% of vehicles Recognition rate enabling possible implementations system in real life.Here they have done detection and counting using deep leaning and have shown the results on screen, with accuracy of 75% .[2]

This work explains how YOLOV3 provides one of the quickest computer identification techniques for objects in a picture with stacked probabilities. Although accuracy varies, speed is the most important component here. A legitimate entity is determined with YOLO, without sacrificing precision, using version 3. YOLOv3 employs DarkNet53, which adds 53 levels to the detection method based, giving us a 106-layer convolutional-based framework for YOLOv3 functionality.Here they have done detection and classification using openCV and results were shown on video screen only , with accuracy of 80% .[3]

To increase performance of the classifier, a concept called as Join Fine-Tuning (JF) was applied. The dataset provided the highest accuracy in 50 and 101 layers of residual network classification and task localisation among numerous cutting-edge techniques, including VGG16, AlexNet, and ResNet50 for recognition.[4].

We have done the Vehicle Detection , Classification and Counting using YOLOv3 . Results shown on screen as well as in excel as table format as shown in Table1.

3 Proposed Method

Using yolo family algorithms for the quickest and most precise multiple item recognition in a frame. Yolo is different from typical networking neural network architecture in that it uses a single convolutional network to forecast circumferencing limits and the estimations that follow from them. [8]It is feasible to stronger competitive advantage the model's end-to-end output since the bounding limits are evaluated in accordance with the probabilities and the model depends on the final weights for their identification.

This allows for the production and analysis of images. Yolo version 3 contains 106 layers overall, with three alternative layer types making up the architecture. The core structure patterns in a remnant structure where activation is easily conveyed to a deeper phase in the neural network and layer 2 together before resizing the image in this example, the open layer, it increases the positioning accuracy, is ranked third.[20] The detecting layer, which carries out detection from three different phases, is second. Furthermore, the outputs of the previous layer are combined with those of the current layer using the concatenation process. In the proposed study, we use the coco dataset, which is simply anticipating the bounding box main objective 05 identifying the class level of confidence. [9]

When non-max suppression is applied, it enables the labelling of many objects in separation masks for image labeling.By ignoring the boundary limit objects acquired in typical settings, key-point finding and segmentation by segments using a panopticon annotations with a total of 81 categories creates one very versatile and powerful and versatile dataset. Using the COCO dataset to identify and classify different types of vehicles in order to gauge the quantity and nature of traffic within a given period.[10]

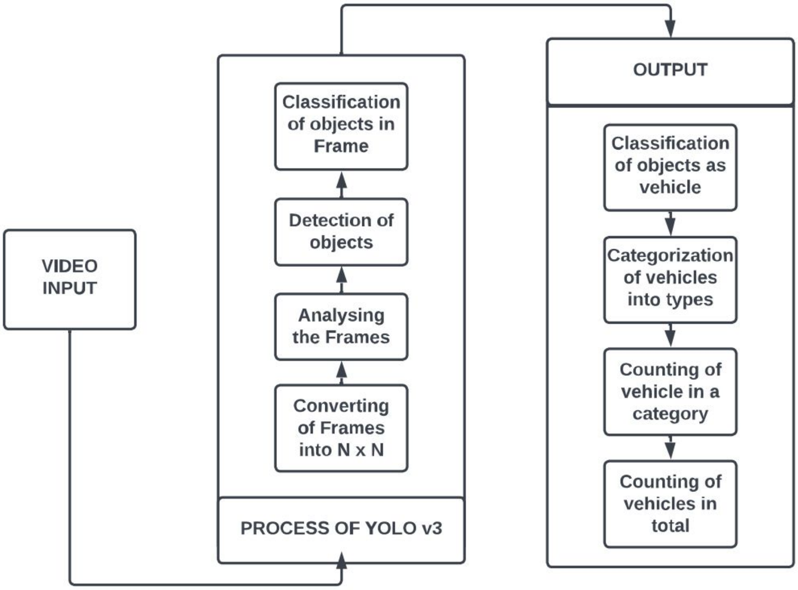

Fig. 2 Methodology of YOLO algorithm

4 IMPLEMENTATION & RESULTS

Starts with Importing a predetermined picture size. Rearrange the image here to desired scale if the size does not satisfy your computer. The program's 32-pitch 416 by 416 input vector can generate a 13 x 13 output frame. During video source, frames are taken into account. Two images are fed into a single end-to-end convolutional neural network, which produces an utilized to distinguish outcome. The output is a collection of region proposals discovered under predefined categories. Every test image has six properties.[11] A coordinate cell has four attributes (x, y, w, h), the feature point (P0) is the fifth attribute, and the stratum point (Sn) seems to be the sixth characteristic.The prediction's centre parameters are x and y, and its length and elevation are w and h. Calculate the probability values following delimiting the boundaries to get the expected groups. [12]

Fig. 3 YOLO Algorithm Working

Non-maximum compression: Select the cell with the highest probability score. Determine the overlap of one box with another box and remove overlapping box with value IOU threshold.

IOU = intersection area/joint area Area of intersection =

$$Areaofintersection = (a2 - a1) * (b2 - b1) \tag{1}$$

Where, a1, a2, b1, b2 are the coordinates of a box.

Link Area = (box1 area + box2 area) - Intersection area

Continue this stage until there squares have a low grade than the specified squares. To establish the object observed in the category, item points and class reliability are employed.[13]

4.1 Dataset & Training

The information is gathered from cameras pointing in distinct angles. The image depicts typical Indian roadways with vehicle kinds such as cars, buses, trucks, motorbikes, pedestrians, and so on. Training: The YoloV3 model's weights are also all pre-trained weights from the COCO dataset. The COCO dataset is a massive collection of tagged picture images. It has 80 class labels for various things.[14] However, the suggested approach basically entails vehicle and passenger segmentation. As a result, it is configured to receive only the required class labels.COCO pre-training weights were utilized and tailored to specific classes. Darknet and yolov3 are pre-trained models for real-time object tracking.

4.2 Working

The following is the training neural network modelling framework. It consists of a 53 layer CNN with three 3 and one 1 convolution layers. Caffe, a well-known deep learning framework, was used for prior and pre-training. Confidence: The presence or absence of an object of any class with a low confidence value indicates that perhaps the networking misread the object. The minimum number for the equal to the coefficient for a moderate recognition filter is 0.5.[15] Threshold: With a default value of 0.3, this represents the maximum null suppress threshold.

For count vehicles using opencv: A collection for detecting and counting the number of vehicles present. On the server, a virtual machine instance is generated for video processing. Traffic footage is collected from cameras facing various directions, and the data is processed in a server-created virtual version. Image extraction is the initial step inside the video processing. This procedure requires detecting variations in an image and using a threshold for detect people and objects.[16] It is commonly conducted on binary pictures upon morphological alteration recovery. The morphological transformation function accepts two variables: the structural element and indeed the reference image. The type of activity is determined by structural elements. The two fundamental actions of morphological alteration are destruction and development.[18] Fig-4 shows the output in a static image. Fig-5 shows the output in a video *dynamically*. Where as the Table-I show the data collected from videos in EXCEL or Spread Sheets in tabular format.[19]

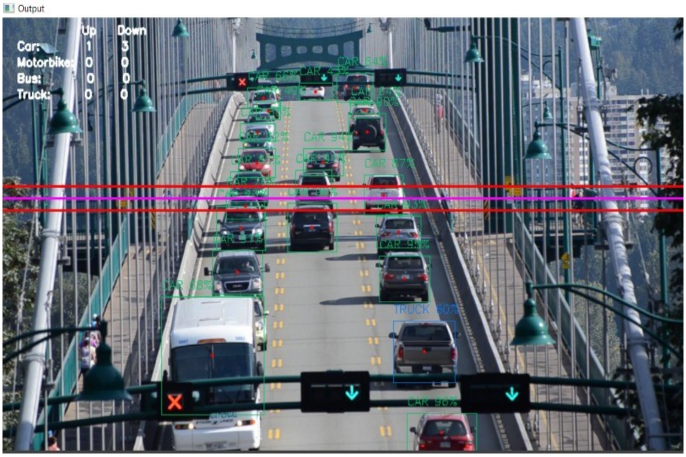

Fig. 4 Output of Classification & Counting (static)

Fig. 5 Output of Classification and Counting & (dynamic)

Table I: Data collected in EXCEL or Spread Sheets

Direction	UP	DOWN
CAR	5	6
TRUCK	3	1
BUS	1	2
TWO WHEELER	2	0

5 CONCLUSION & FUTURE SCOPE

Leveraging YoloV3, we suggested a vehicle recognition, categorisation, and enumeration system that classifies small and large cars into different groups using a convolutional neural network. The technology for image processing enhances the vehicle monitoring system in this technology, and the object tracking system counts and categorises cars according to the input load and category recognition criteria.

The COCO dataset's which was before parameters are utilised to initialise all weights in the YoloV3 algorithm. The COCO dataset contains a large number of labelled image objects. A traffic surveillance approach using YOLOv3 has now been constructed without observing the movements of the vehicles. Tallying is done by simply measuring the distance between the vehicle's centroid and the boundary line.Finally, this approach saved the data in the shape of type of vehicle and number of vehicles count inside a CSV file.

Vehicle detection,classification and counting is implemented by using deep learning technology YOLO. For better understanding and extraction of every vehicle we can develop this project as application programming interface and we can maintain the data in server / cloud with every vehicle future and its accuracy.

References

[1] Moving Vehicle Detection and Classification Using Gaussian Mixture Model and Ensemble Deep Learning Technique (hindawi.com)

[2] https://ieeexplore.ieee.org/document/6908506

[3] https://ieeexplore.ieee.org/document/9194678

[4] Manana Mduduzi, Tu Chunling and Owolawi Pius Adewale, "A Survey on Vehicle Detection Based on Convolution Neural Networks", 3rd IEEE International Conference on Computer and Communications, 2017.

[5] Jung Heechul and Jihun Min-Kook Choi, "ResNet-based vehicle classification and Localization in traffic surveillance systems", IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017.

[6] Yong Tang, Zhang Congzhe, Gu Renshu, Li Peng and Bin Yang, Vehicle detection and recognition for intelligent traffic surveillance system, New York:Springer Science+Business Media, January 2015.

[7] Hadi, R. A., Sulong, G., & George, L. E. (2014). Vehicle detection and tracking techniques: a concise review. arXiv preprint arXiv:1410.5894.

[8] B Bilgic, BK Horn and I Masaki, ”Efficient integral image computation on the GPU[C]”, Intelligent Vehicles Symposium (IV), pp. 528-533, 2010.

[9] S Fazli, S Mohammadi and M Rahmani, ”Neural network based vehicle classification for intelligent traffic control[J]”, Int J Softw Eng App, vol. 13, no. 3, pp. 17-22, 2012.

[10] Redmon Joseph and Ali Farhadi, ”YOLOV3: In Incremetal improvement”, Cornell University-Computer Vision and pattern recognition, April 2018.

[11] Redmon Joseph, Santosh Divvalay, Ross Girshick and Ali Farhadi,You Only Look Once Unified, Real-Time Object Detection AI research, May 2016.

[12] Guerrero-Gomez-Olmedo, R., Torre-Jimenez, B., Lopez-Sastre, R., Maldonado-Bascon, S., Ooro-Rubio, D. (2015). Extremely overlapping vehicle counting. In Iberian Conference on Pattern Recognition Image Analysis. Springer International Publishing, (pp. 423–431)

[13] S. Yu, Y. Wu, W. Li, Z. Song, and W. Zeng, “A model for fine-grained vehicle classification based on deep learning,” Neurocomputing, vol. 257, pp. 97–103, 2017.

[14] F. C. Soon, H. Y. Khaw, J. H. Chuah, and J. Kanesan, “Semisupervised PCA convolutional network for vehicle type classification,” IEEE Transactions on Vehicular Technology, vol. 69, no. 8, pp. 8267–8277, 2020.

[15] V.MuniSekhar, K.V.G.Rao,K.Sambasiva Rao, M Gopichand “Comparing the Capacity,NCC and Fidelity of Various quantization Intervalson DWT”, International Conference on Innovations in Computer science Engineering(ICICSE-2015), published in Journal of Advances in Intelligence Systems and Computing, Springer ISBN 978-981-10-0417-9, August 2015.

[16] A Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar and L. Fei-Fei, ”Large-scale video classification with convolutional neural networks”, Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2014.

[17] Rublee, E., Rabaud, V., Konolige, K., Bradski, G.R. (2011). Orb: an efficient alternative to sift or surf. In 2011 International Conference on Computer Vision. https://doi.org/10.1109/iccv.2011.6126544. IEEE.

[18] Vehicle Classification in Traffic Surveillance System using YOLOv3 Model

[19] P. Avinash, Ranjith Reddy S, M. Kavitha “Analysis On Market Prediction Using machine Learning And Deep Learning Techniques”, JOURNAL OF APPLIED SCIENCE AND COMPUTATIONS, Feb-2022,ISSN NO : 1076-51341

[20] Vehicle Classification and Counting System Using YOLO Object Detection Technology

[21] A Simple Vehicle Counting System Using Deep Learning with YOLOv3 Model

[22] Prem Kumar Bhaskar and Suet-Peng Yong, ”Image Processing Based Vehicle Detection and Tracking Method”, IEEE, 2014.