ICCIP-2019
Skilled Engineers for Society

International Conference on Communication and Information Processing (ICCIP-2019)
Available on: Elsevier-SSRN

Freight Analysis Using YOLOv2

Snehal Kadam, Akash Hatalge, Abhishek Balip, Avinash Powar

Rajarambapu Institute of Technology, Islampur 415414, India

Abstract

Monitoring traffic of India and calculating the peak hours and density count in a single day helps to develop a required travel and traffic volume estimates, which is required for satisfying all the needs in the planning of roads, its construction, its maintenance and overall administration of the state. Vehicle counting is an important aspect to understand the traffic load and optimize the traffic signals. Detection of vehicles is expected to be more efficient and robust in number of sceneries. Due to improvement in various algorithms and research work, detection mechanism of traffic data analysis has made a significant improvement over traditional methods. Traditional machine learning algorithms and computer vision for object detection now running under slow response time. This problem can be solved by modern architectures and algorithms based on ANN (Artificial Neural Network), like YOLO (You Only Look Once) without any major losses. YOLO and its versions achieved a jaw-dropping performance in computer vision and had achieved a great success in object detection and classification. In this paper, we are presenting vehicle counting, detection and classification based on YOLOv2. Some video sequences have been taken and tested with the planned algorithm. The results can be a solution for planning of new roads or any other diversions for heavy vehicles can be considered during the peak time. A detection mechanism through YOLOv2 differs from other roadway sensors, such as radar or inductive loops, which provide data only regarding traffic flow and density, and do not provide information about the type of the vehicle in real time.

Keywords- Convolutional Neural Network, YOLOv2 Algorithm, Vehicle Counting, Vehicle Detection

1. Introduction

One of the main requirement is that all the vehicles should be able to recognize and understand their surroundings for better traffic management, faster and appropriate traffic flow and for faster goods transportation. In classical method, by using various sensor types like ultrasonic devices, radars, cameras, etc., all the vehicles collect information from their surroundings. Comparing to other sensor types, cameras give the more detail information about vehicle's surrounding in terms of texture information and high resolution. Understanding and interpretation of visual information taken by the cameras still shows very challenging and difficult task for especially computers when there is a need to perform it in real time, Aleksa Corovic and Velibor Ilic, 2018. The fluctuating weather, lighting conditions and miscellaneous surroundings could make this task even more complex. To solve this problem, suitable solutions are already developed for traffic detection in automotive industry. There are some object detection algorithms like ViNotion, StradVision, NVidia, Mobileye afore-mentioned and the results generated by them are used by big companies at that time. However, all of these algorithms are closed and now there is not much information available about them. Unlike these algorithms, we are using different one. On the other hand, there are some research papers has

been available for detection of objects. Few of older papers uses algorithms such as SVM or sliding window method and newer approaches uses deep CNN, like Fast R-CNN or Faster R-CNN, but they gives the slower response time compared to YOLO object detector. Due to CNN implemented in YOLO platforms, various objects can be detected, classified and tracked from video. Most of the older research papers have focused on object detection in general and not on the traffic participant's detection. The YOLO based on CNN gives the outstanding results in various object detection sceneries due to which we can achieve the real time performance rate. In this paper, we are presenting YOLOv2 algorithm which is able to detect objects accurately compared to YOLO and all other older algorithms.

2. Objectives

To present the architecture and working of YOLO and YOLOv2, the selection parameters which are required for detection and classification of various objects is the first objective, Matija Radovic and Offei Adarkwa, 2017. To demonstrate one successful application of YOLOv2 on real time object classification and detection from video is the second objective. The third objective is based on the results of the detection, classification and counting different vehicles, suggest possible ways for the faster and appropriate goods transportation.

3. Background and related work

3.1. Detection based approaches

Before the rise of CNN in computer vision, many existing object detection algorithms adopt sliding window or SVM and based on that they used hand crafted feature techniques or HOG features and searching to recognize and localize objects in an image. However, all the detected vehicles are considered as vehicles only due to shortage of high level information in terms of various vehicle types, Jiasong Zhu and Ke Sun, 2018. Since size, capacity and shape of various vehicles have various effects on traffic load, hence recognizing various vehicle types in traffic is much essential. Thereafter, some people put more classifiers to detect various vehicle types, but they require more cost for parameter optimization and computation.

3.2. Movement based approaches

Many existing systems tried to estimate the traffic density using movement based tracking techniques such as background subtraction and optical flow. But these methods could only work well on simple traffic sceneries like roads in the rural areas and expressway. They fail in the urban sceneries due to distraction in the surrounding and complex local ground conditions. Also, there are few vehicles which are present in the frame, but still they are not accurately estimated.

In short, motion and detection based approaches are very sensitive to quality of the video and the road conditions. Also, many existing methods could not give the accurate number of various vehicle types.

Thereafter, artificial neural network has been used. In that, lots of working on deep neural network (DNN) and convolutional neural network (CNN) has been done, LI Suhao and LIN Jinzhao, 2018. In one of the research work, urban traffic density estimation has been done by using advanced deep neural network. They had achieved accuracy and efficiency too. But still, we can get more accuracy, efficient detection of vehicles and tracking algorithm while achieving high processing speed.

After 2013, CNN become the standard for detection of objects. There are numerous applications in the civil infrastructure where CNN is used in one of the research paper. Traffic surveying, power line surveillance, disaster management are some of the examples, Matija Radovic and Offei Adarkwa, 2017.

It is essential to increase the efficiency, accuracy and safety of the device. They worked on aerial images set for automated and efficient object identification. CNN can classify and detect various objects with high level of computational efficiency and accuracy of about 97.5%. Thereafter, CNN implemented in YOLO platform, so then objects are able to detect, classify and track from the video in real time.

YOLO which uses the deep learning and CNN for object detection, stands out from its "competitors" because, as the name itself shows it only needs to "look" each image only once. YOLO has proven to be more efficient algorithm compared to traditional machine learning algorithms. It can detect and classified various objects. By using YOLO, research has been done to detect the traffic signs to solve the problems in

autonomous driving. They said, unlike other region proposal CNN architectures, YOLO uses global features instead of local features to predict the bounding boxes and labels. For detecting traffic signs, they mainly kept their focus on analyzing costs and benefits of using YOLO. YOLO does fast object detection, but it faces difficulty while detecting spatially small objects. YOLO works on entire image instead of only within the bounding boxes. Thus, while working on PASCAL VOC dataset, it is much informative, but it is much difficult to detect traffic signs as they appear in various places.

Thereafter, in one of the research work, they focused on Chinese traffic signs detection and differentiate it with other countries traffic signs. They had been used deep convolutional network. They had used end-to-end convolutional network, Jianming Zhang and Manting Huang, 2017. To detect small traffic signs, they had divided input images into dense grids and obtain finer feature maps. By working on this, they had achieved more robust and faster result. The detection speed accomplished by them was 0.017 sec.

YOLOv2 is an enhanced version of YOLO. YOLO9000 is build on top of YOLOv2, but can be trained with the dataset combining top 9000 classes from ImageNet and COCO dataset. Still single shot detector (SSD) and faster R-CNN running significantly faster, but in one research work, they proposed a system to jointly train on object classification and detection. By using this system, they had trained YOLO9000 on COCO and ImageNet dataset. YOLO9000 is robust and it fills the gap of dataset size between detection and classification. Their used representation of WordTree gives richer and more detailed output for classification of image, Joseph Redmon and Ali Farhadi, 2016.

Furthermore, YOLOv3 came into the market and research work has been done onto that. They had worked on traffic participants detection. For autonomous cars, they had used the newest YOLOv3 algorithm since object detection would be the key for software component in the upcoming generation. They had trained the network for mainly 5 objects (lights, traffic signs, truck, pedestrian and car). They had represented their effectiveness in various climate conditions such as snow and night, fog, bright and overcast sky. YOLO can solve this problem, but they had mainly focused on the response time and so had used YOLOv3. They had worked on common objects in context (COCO) dataset. By the results proposed by them, they had achieved the specializing YOLO network for real time multiple objects detection and tracking in traffic participants, Aleksa Corovic and Velibor Ilic, 2018. They considered YOLO algorithm as a strong base for the detection of objects.

4. Methodology Implemented

4.1. Existence and architecture of YOLO

Object detection is one of the integral part in computer vision. So, in older times, CNN was used for that, but then that has been overcome by R-CNN (Recurrent Convolutional Neural Network). Since, R-CNN was working on only 2000 regions from the image called region proposals, so instead of trying huge number of images, this is the better way. 2000 region proposals are generated by selective search algorithm. But still they require a huge amount of time to train the network. It takes 47 seconds for each individual test image. Also, as the selective search algorithm is fixed one, so learning could not be happen, so ultimately it is the bad way. Thereafter, Fast R-CNN came. It is faster than the R-CNN because it doesn't work on 2000 region proposals to the CNN every time. Instead, it uses only single image once and a feature map is created from it. But though it is good algorithm, it has overcame by faster R-CNN, since fast R-CNN and R-CNN works with the selective

search algorithm and since it is time consuming and slow process, it affects onto the performance, Venceslav Kafedziski and Sinisha Pecov, 2018. Faster R-CNN is analogous to the Fast R-CNN, but instead of using selective search algorithm, it uses a separate network to predict the region proposals. But as all above algorithms only uses and works on the parts of the images to localize the object within the image, they are not considering a whole image to process. So, now, YOLO has been arrived. YOLO predicts the bounding boxes and the class probabilities for these boxes. It takes an image and divides into S*S grids. If the center coordinate of an image falls into a grid, the grid is responsible for detecting the objects, Yu Zhao and Quan Chen, 2019. However, YOLO predicts the bounding boxes of objects in all grids, the probability vectors of all classes and the location reliability at the same time, hence solves problem of one-shot, Junyan Lu and Chi Ma, 2018. The network outputs a class probability and offset values for each bounding box. The class probability having the above threshold value is selected and used to locate the object within the image. YOLO having the faster magnitude (45 frames per second) uses the 1*1 convolutional layer + 3*3 convolutional layer instead of inception module, Ying Liu and Zhishan Ge, 2018. YOLO network structure consists of 24 convolution layers and 2 full connection layers, as shown in the YOLO architecture, (Fig.1.), Ying Zheng and Hong Bao, 2018.

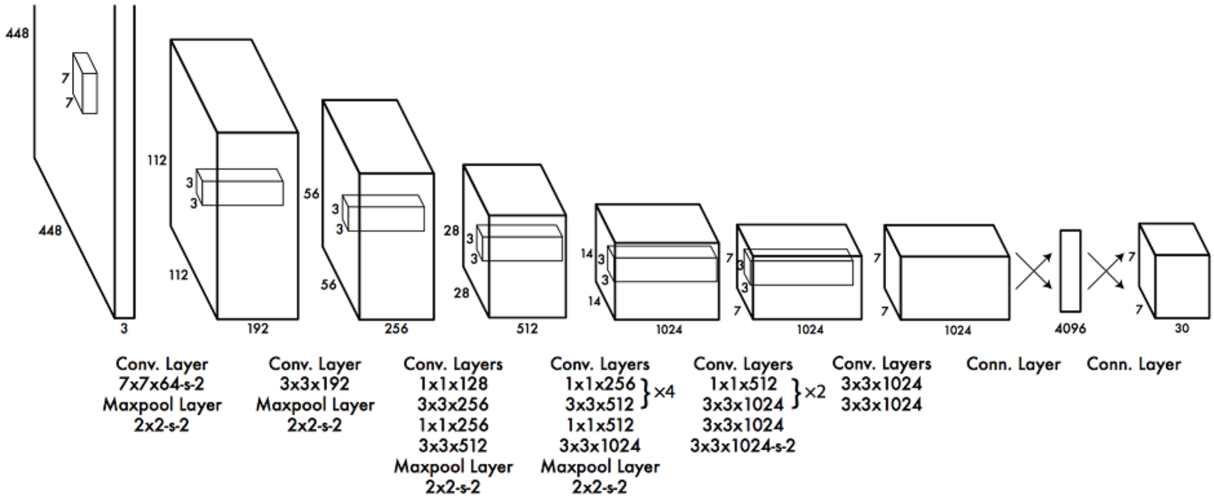


Fig. 1.YOLO architecture

But YOLO struggles for the small objects within the image, like flock of birds. So, this is the limitation of the YOLO. So, then YOLOv2 has been arrived. YOLOv2 uses the idea of faster R-CNN instead of using full connection layer to predict the coordinates of bounding boxes as in YOLO. It adds the anchor boxes which results into effectively improvement in the recall rate. Also, instead of working with the fixed image size as input like in YOLO, YOLOv2 adjusts the input image size so that it has good detection effect on the multi-scale input images.

4.2. Our work

We have taken few video samples that are available on internet to test the YOLOv2 algorithm and find out the results. So, first of all, we have installed python 3.5.1 and anaconda. You can use any editor which supports python files like spyder. Then we have installed openCV and downloaded the darkflow repository. Since we are using and working on YOLOv2, we have downloaded the YOLOv2 pre-trained weights. Then, we had used one video as input and store results in second video after various object detection and classification. Since we are working for different types of vehicles only, so our main focus is on cars, trucks and bikes. Here, YOLO only displays objects detected with a confidence of 0.25 or higher than that. We can change this threshold value as well. After detection and classification of these vehicles, counting of respective vehicle by their type is done. For counting, we have drawn two lines in video and named them as “ROI” (Region of Intersection). By the coordinates of this ROI and the labels of the respective objects, we have counted different types of vehicles separately. Thus, traffic analysis can be easily done. That is, as our main aim is to analyze the traffic caused by the heavy vehicles such as truck, so it can be easily detected and analyzed. Due to this fine & crystal clear analysis, we can easily suggest possible solution to overcome this problem. That is the result can be a solution for planning of new roads or any diversions for heavy vehicles can be considered during the peak time. Hence, this gives the advantages like low transportation time will be then require for all the vehicles, faster goods transportation will be then done and obviously it results into the reduction of the traffic. Our flow chart looks like:

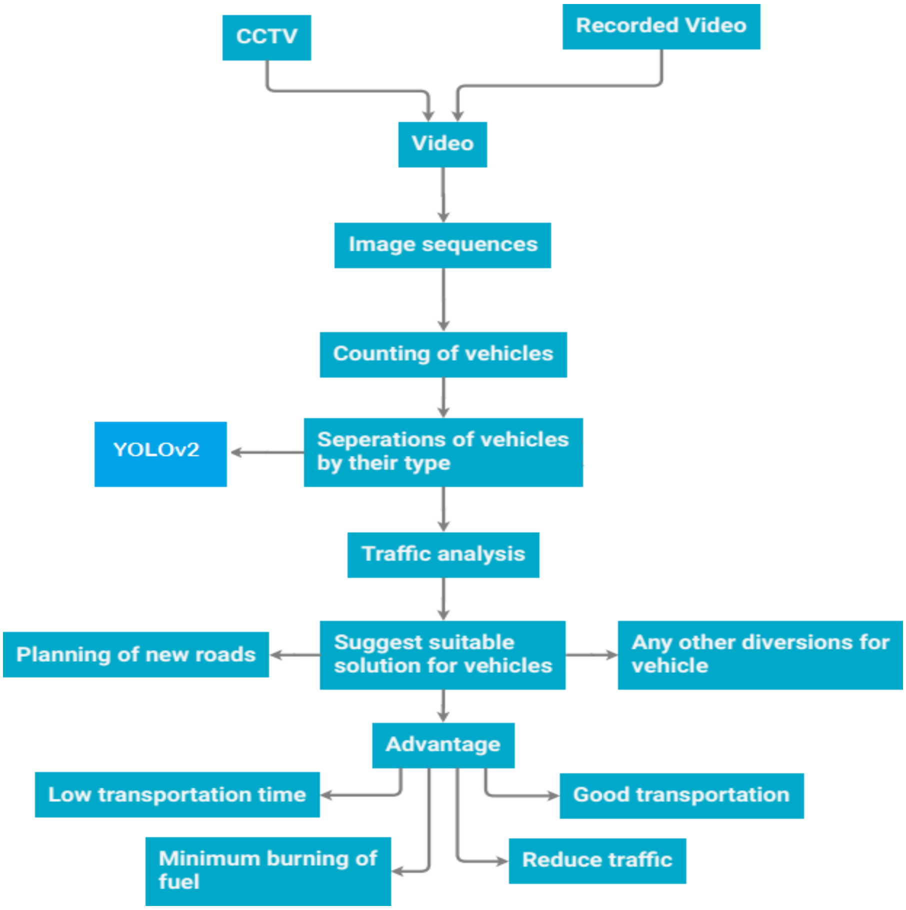


Fig.2. Project flow chart

4.3. Results

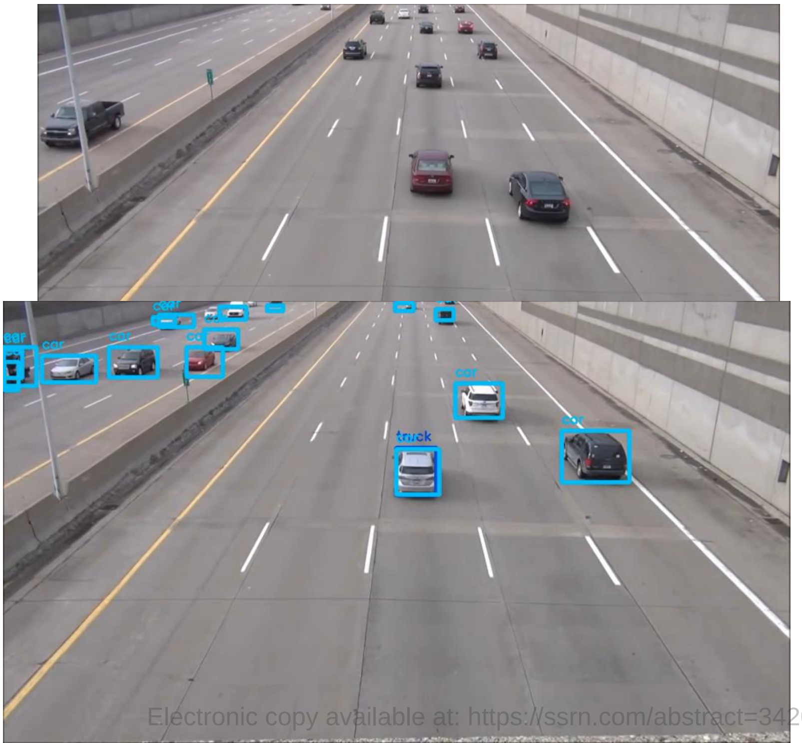
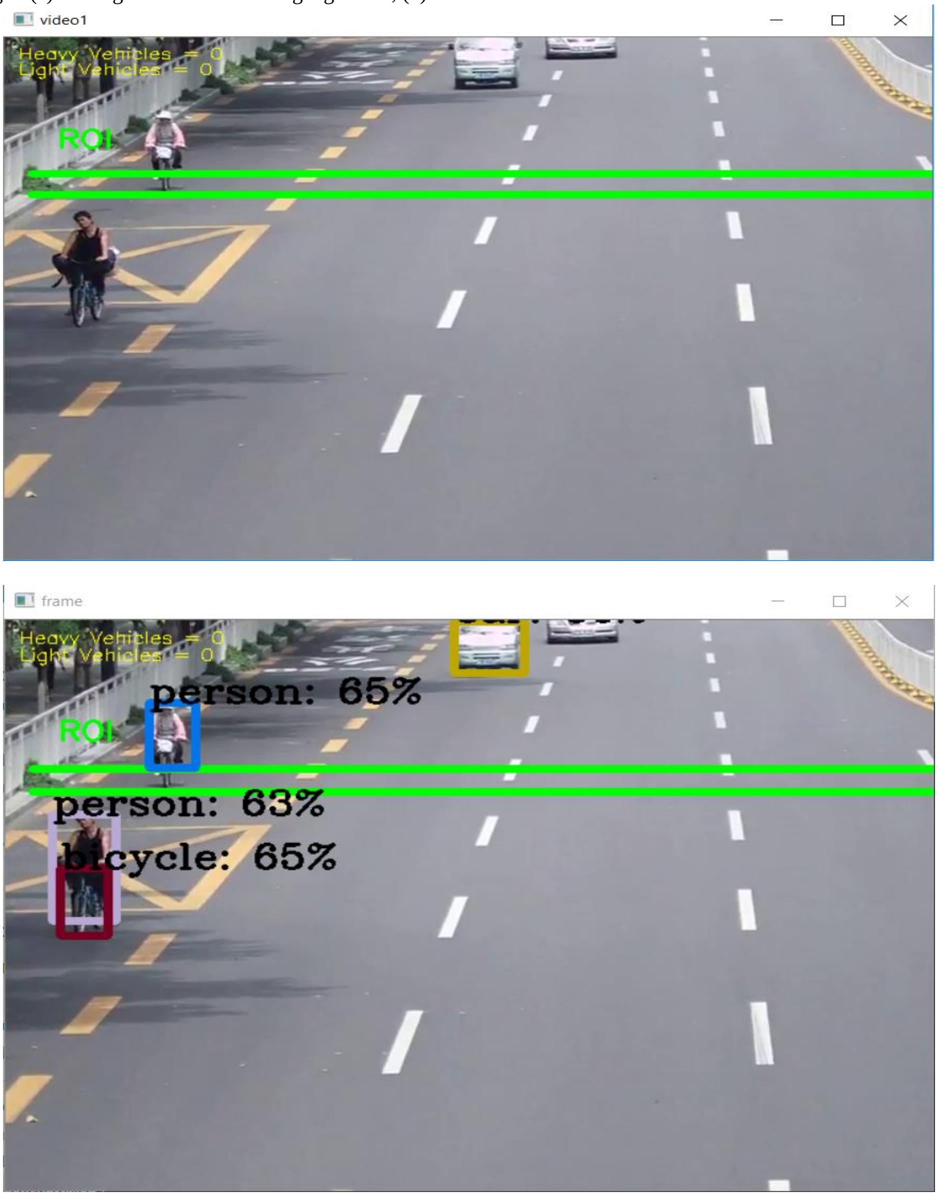


Fig.3. (a) Testing video 1 before using algorithm, (b) Resultant video shows detection and classification of vehicles



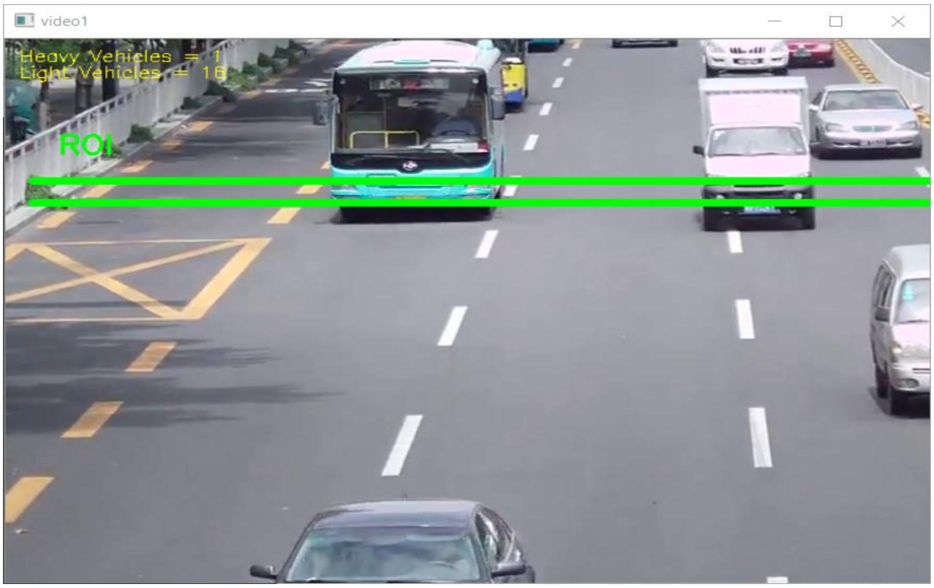


Fig.4. (a) Testing Video 2, (b) Actual processing of input Video 2, (c) Resultant video shows the counting of heavy and light vehicles separately.

Table 1. Comparison Table for object detection and classification

Publication-Year	Title	Framework	Accuracy (mAP)
IEEE Conference on Computer Vision and Pattern Recognition (CVPR)-2016	You Only Look Once: Unified, Real-Time Object Detection	YOLO	63.4%
Journal of Imaging-2017	Object Recognition in Aerial Images Using Convolutional Neural Network	CNN	97.8%
IEEE Conference on Computer Vision and Pattern Recognition (CVPR)-2017	YOLO9000: Better, Faster, Stronger	YOLOv2, YOLO9000	78.6%

Journal of Computer and Communications-2018	A Vehicle Detection Method for Aerial Image Based on YOLO	YOLO	76.7%
ELSEVIER-2018	Object detection based on deep learning for urine sediment examination	DFPN	86.9%
ICCIP-2019, Available on: Elsevier-SSRN	Freight Analysis Using YOLOv2	YOLOv2	98%

Table 2. Comparison Table for counting of vehicles

Publication-Year	Title	Framework	Accuracy (mAP)
Conference on Technologies and Applications of Artificial Intelligence (TAAI)-2018	A YOLO-based Traffic Counting System	YOLO	100%
ICCIP-2019, Available on: Elsevier-SSRN	Freight Analysis Using YOLOv2	YOLOv2	99%

5. Conclusion

YOLOv2 is faster than other detection algorithms over a diversity of detection datasets. It is able to run at various size of images and can provide a smooth tradeoff between accuracy and speed in real time processing, Joseph Redmon and Ali Farhadi, 2016. We introduced and used YOLOv2. Since YOLOv2 is much better to detect and classify various objects, different types of vehicles can be detected, classified and labeled. Due to this, there is ease in management of traffic as well as it is better for the goods transportation by heavy load vehicles such as truck, buses, etc. Classification with counting of different vehicles separately gives awesome result in traffic density estimation. By the result of YOLOv2, we can easily suggest to the government for the construction of new roads for heavy load vehicles or any other diversions for heavy load vehicles can be suggested so that other vehicles will not be get stuck more time due to those vehicles in the traffic and can get low transportation or travelling time. As the YOLOv2 can detect spatially small objects too, so each and every object in that respective image or video can be easily detected, counted and classified. Hence, it gives the jaw-dropping performance in all ways.

As part of our future work, we would like to improve our detection results using even more powerful matching methods for assigning weak labels. We would like to switch and work on the different architecture to some more developed framework.

6. References

Aleksa Corovic, Velibor Ilic, Sinisa Duric, Malisa Marijan, and Bogdan Pavkovic, "The Real-Time Detection of Traffic Participants Using YOLO Algorithm", Serbia, Belgrade, November 20-21, 2018., 26th Telecommunications forum TELFOR 2018.

Matija Radovic, Offei Adarkwa, Qiaosong Wang, "Object Recognition in Aerial Images Using Convolutional Neural Networks", Article in Journal of Imaging, 14 June 2017.

Jiasong Zhu, Ke Sun, Sen Jia, "Urban Traffic Density Estimation Based on Ultrahigh-Resolution UAV Video and Deep Neural Network", IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 11, NO. 12, DECEMBER 2018, p. 4968.

LI Suhao, LIN Jinzhao, LI Guoquan, BAI Tong, WANG Huiqian, PANG Yu, "Vehicle type detection based on deep learning in traffic scene", 8th International Congress of Information and Communication Technology (ICICT-2018).

Jianming Zhang, Manting Huang, Xiaokang Jin, Xudong Li, "A Real-Time Chinese Traffic Sign Detection algorithm Based on Modified YOLOv2", Article in algorithms MDPI, 16 November 2017.

Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

Joseph Redmon, Ali Farhadi, “YOLO9000: Better, Faster, Stronger”, 2017.

Venceslav Kafedziski, Sinisha Pecov, Dimitar Tanevski, “Detection and Classification of Land Mines from Ground Penetrating Radar Data Using Faster R-CNN”, 2018 26th Telecommunications Forum (TELFOR), 2018.

Yu Zhao, Quan Chen, Wengang Cao, Jie Yang, Guan Gui, “Deep Learning for Risk Detection and Trajectory Tracking at Construction Sites”, IEEE Access, 2019.

Junyan Lu, Chi Ma, Li Li, Xiaoyan Xing, Yong Zhang, Zhigang Wang, Jiuwei Xu, “A Vehicle Detection Method for Aerial Image Based on YOLO”, Scientific Researcher An Academic Publisher, Vol.6 No.11, November 2018.

Ying Liu, Zhishan Ge, Guoyun Lv, Shikai Wang, “Research on Automatic Garbage Detection System Based on Deep Learning and Narrowband Internet of Things”, Journal of Physics: Conference Series, 2018.

Ying Zheng, Hong Bao, Xinkai Xu, Nan Ma, Jialei Zhao, Dawei Luo, “A Method of Detect Traffic Police in Complex Scenes”, 2018 14th International and Security (CIS), 2018.

Yixiong Liang, Zhihong Tang, Meng Yan, Jianfeng Liu, “Object detection based on deep learning for urine sediment examination”, ELSEVIER-Biocybernetics and biomedical engineering 38 (2018) 661-670.

Jia-Ping Lin and Min-Te Sun, “A YOLO-based Traffic Counting System”, Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2018.