International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-2, December, 2019

IoT with Cloud Centric Vehicle Detection and Counting System for Smart Traffic Surveillance

S. Saravanan, K. Venkatachalapathy

Abstract*: *The efficient management of road traffic is one primary facet of many, in smart cities. Traffic overcrowding can be managed successfully, if prior estimation of the number of vehicles that will pass though a crowded junction in a specific time is known. This paper introduces a methodology which targets vehicle extraction on videos covering vehicles. To resolve the problem of current vehicle detection such as the need of detection accuracy and slow speed, an improved YOLOv3 vehicle detection is utilized. The k-means clustering used to group the bounding box around the vehicle in training dataset. The method for calculation of loss with respect to the length and width of the bounding boxes was recovered through the implementation of the batch normalization process. Finally, to improve the feature extraction of the network the high repeated convolution layer are removed. The experiment results are carried out on the BIT-vehicle validation datasets which shows the improvement of mean Average Precision (mAP) could certainly reach 95.6%.

Keywords*: *Traffic Surveillance, YOLOv3, k-means Clustering, accuracy, IoT.

I. INTRODUCTION

Intelligent Transportation System (ITS) plays a crucial role in enhancing vehicle to vehicle communication and also between the vehicles to infrastructure. Intelligent Transportation System utilizes advanced information and communication in order to decrease traffic congestion and reduce accidence. Use of IoT for automated traffic light control system is presented. IoT enabled smart cameras and infrared (IR) sensors are proposed to automate the traffic signals. Wifi transmission of the observed data from IR sensor is sent to the raspberry-pi controller (ruling out the unnecessary waiting for vehicles in most crowded region). IOT based traffic control system makes use of the software (image processing) and WiFi communication module (sending details related to vehicle count) along with the raspberry-pi and Arduino as the hardware components.

The proposed traffic control system provides a lot of benefits. It reduces unwanted delays and grid locks. It also aides in detection of accidents, vehicle parked violating the earmarked parking zones, obtain traffic flow data's, and compliance monitoring of traffic. This work is aims at providing solution to pertinent traffic control problems to detect and count the number of passing vehicles.

The Control Centre obtains video log of the real time traffic is obtained from a stationary video camera installed above the road. The process involves extensive image processing due to the limitation of computing capabilities of cameras. It is important to get the right kind of picture as there will be distortions in scene illumination due to factors like shadows from buildings, moving shadows and glare caused by sun, weather condition and various angles of cameras. The paper is segregated into five sections. An introduction to the proposed work is given in Section I. Section II deals with the literature cited on related works. Methodology proposed for this work is elaborated in Section III. Section IV speaks on the results and the same is discussed in detail. The Inferences of the work and concluding remarks are deliberated in Section V.

II. RELATED WORKS

In order to solve and overcome the existing disadvantages in urban traffic problems (lack of vehicle information, low accuracy of vehicle information retrieval) intelligent transportation was developed. Two categories of common vehicle detection methods are machine learning (ML) and deep learning (DL) models.

The classical ML employs the Histogram of Oriented Gradient (HOG) model [1]. This helps in the extraction and classification of the features of the images using SVM (Support Vector Machine). A deformable part model (DPM) is used to detect and obtain good result [2]. These methods are very complex, need human involvement and take more time making it impractical for applications.

DL becomes more familiar research direction in the recent past. Detection and recognition methods of objects exhibit better performance when compared to traditional methods. Methodologies using Convolutional Neural Network were proposed for the vehicle detection in order to procure richer vehicle features. It uses only the vehicle images in order for network training and obtains the vehicle type features automatically. Unsupervised method using network pre-trained (sparse code) and vehicle classification has been done using softmax. By the use of R-CNN, selective search technique is applied for extracting around 2000 regions from the image referred as region proposals. To improve CNN based image classification methods as implemented in SPP-net, fast R-CNN, faster R-CNN [3] and R-FCN. Use of Adaptive neural network was presented which helps in extraction of features of different scales by dividing the last layer into several networks. An improved CNN, MS-CNN, has been proposed to be used in detecting vehicles by dividing the regions, thus improving the accuracy, memory and computation [4]. Use of R-CNN was proposed to achieve improved vehicle detection.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author

S. Saravanan*, Assistant Professor/Programmer Department of Computer and Information Science, Annamalai University, Chidambaram, India. Email: aucissaran@gmail.com

Dr. K. Venkatachalapathy, Professor, Department of Computer and Information Science, Annamalai University, Chidambaram, India. Email: omsumeetha@rediffmail.com

Retrieval Number: B4255129219/2019©BEIESP
DOI: 10.35940/ijeat.B4255.129219
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Combined use of Faster R-CNN, VGG16 and ResNet-152 was used for vehicle detection for good accuracy, slow speed but was found to be impractical for the real-time implications. Redmon et al. converted direct object detection to regression and proposed dedicated solution for object recognition method using YOLO, for improvising the speed and accuracy [5]. Redmon et al. proposed YOLOv2 which greatly improved the speed of object detection while keeping the detection accuracy [6]. K-means ++, a clustering algorithm was proposed which selects 6 anchor boxes with different sizes in training dataset. YOLOv2 was used for improvising feature extraction thereby adopting multi-layer fusion strategy and removing repeated convolution layer.

A traffic control model is presented that makes the main server to obtain the images transmitted by the wireless transmitter and based on the traffic density the time span of red light was determined [7]. Method to control traffic congestion - continuous infrared by using sensor sources (the intensity of light reflected back to the sensor we can detect the existence and non-occurrence of the vehicle), Drawback – output based on temperature. IOT based traffic management system [8] uses embedded circuit that works on RFID with clustered system. The data regarding the detection of the vehicle location is obtained by employing sensors and RFID and the same is sent to centralized controlling center for further processing [9]. Improved parking management [10] - Optical Character Recognition (image processing) for verifying employee and extracted number plate forwarded to server system ,only registered candidate allotted for parking. A lane base traffic monitoring system [11]- ultrasonic senses (sensed data processed using controller and transferred to server using WIFI module). Automatic parsing of emergency vehicle is carried out using a raspberry-pi, Node MCU and RFID. Based on the acquired data, the reader controls the traffic signals by using sensors in vehicles [12]. A scalable architecture is proposed for urban IoT-Lambda Architecture by absorbing the data through the web services and storing the data. The use of cloud-computing and big-data management in decision making in traffic management system in smart city concept is studied [13].

III. PROPOSED METHODOLOGY

One of the arduous problems encountered is to optimize traffic flow on roads to bring in efficiency in the transport management system. The use of traffic analysis comes in handy for solving real time problems in and for obtaining statistics of traffic density. Analyzing the traffic density statistics data allows us to suggest real methods to increase traffic safety. The automated detection and counting of vehicles from the video done by using improved YOLOv3 technique. The real time video recording of the traffic flow stream is picked by the cameras mounted on the road sides. Once the video is acquired, the system detects and counts the vehicles using YOLOv3 network. Finally, once the vehicle count is obtained, the system communicates the same to central control system. The detailed work is explained in the following sub-sections:

A. Vehicle Detection and Count

Redmon et al. proposed the YOLO object detection method. As illustrated in Figure 1, YOLO splits the image into S*S grids and foresees the bounding box and class probability for every individual cell grid. Each bounding box consists of five values. The width and height of the box with respect to the image is given by w and h. The coordinates of the box relative to the grid cell is denoted by x and y respectively. The confidence score represents the reliability of existing object in the box, which defined as.

$$\textit{Confidence score} = \textit{Probability (object) X IOU truth Prediction} \quad (1)$$

Equation (1) represents the probability of the object falling into the grid cell. Intersection over Union (IOU) shows both the real box as well as predicted bounding box. Then, the bounding boxes with less confidence score with low threshold values are removed. Finally, the non-Maximum suppression (NMS) method is implied to evict the redundant bounding boxes.

Figure 1. YOLO Object Detection

A new version of YOLO, namely, YOLOv2 with a new Darknet-19 architecture was proposed to increase the YOLO likelihood accuracy, which works on the removal of the fully connected and batch normalization at each layer. In order to obtain the boxes, the anchor model of Faster R-CNN and K-means clustering is used. Further, bounding box were re-oriented with direct prediction. It is found that the speed and accurateness of object recognition is significantly enhanced with YOLOv2 when compared with YOLO.

When there are a plethora of classes to be identified and the difference among the classes are huge, in those case and similar, the YOLO v2 can be applied. Similarly, in the detection of smaller objects, YOLO v2 has improved precision. To detect vehicles, this paper uses an improved YOLOv3 framework to increase the vehicle detection speed with a smaller number of frames. Clustering analysis was applied on the size of the bounding boxes in the BIT-Vehicle training dataset.

The size of the anchor boxes for vehicle detection were selected by implementing distance function in YOLOv3, rather than traditional Euclidean distance. As shown in Equation (2), the IOU was implemented as evaluation metric.

$$d(box, centroid) = 1 - IOU(box, centroid) \quad (2)$$

For vehicle detection, the size of the anchor box which is appropriate could be discovered by undergoing the clustering analysis with the k- means on the training datasets.

While the YOLOv3 conventionally uses only 7 convolutional layers and 6 pooling layers, the improved YOLOv3 method uses additionally three convolutional layers, which aides in extraction of features of the vehicle features with enhanced accuracy. The computation resources increase with a rise in the depth of the network. So, to avoid the increase in computation resources 1x1 convolutional layer is added which significantly saves memory resources. The improved YOLOv3 network architecture is illustrated in Figure 2.

The vehicle images are obtained from the surveillance cameras when it passes through the road. So, when the vehicle is far from the camera, it appears smaller in the image. When it is near to the camera, it takes up large space in the image. Therefore, even if the vehicle type is identical, the size may be different in the image. In order to enhance the loss function normalization, the training of the dataset is done using YOLOv3. The overall loss function makes use of the sum of squared error as the loss function for localized prediction. The root number error serves as the loss function for width and height while the SSE serves as the loss function for confidence. In error loss function is determined by the use of binary cross entropy in the case of YOLOv3IOU.

Once the total vehicles count is obtained from the video frames at specified interval time, the traffic density value is to be communicated via internet. This information paves way to estimate the density of traffic in a specific time interval and the same is transferred to the central server system from the raspberry-pi.

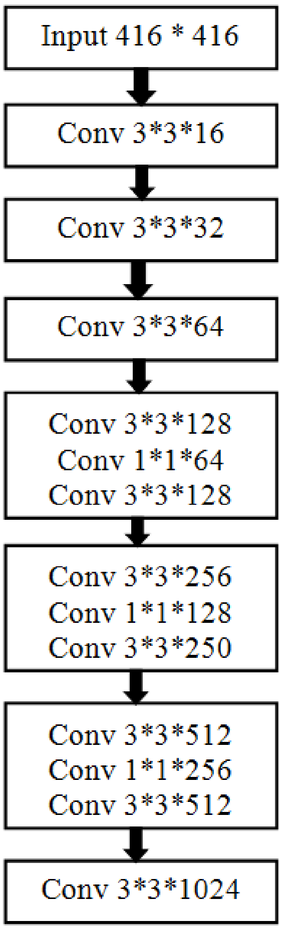

Figure 2. Improved YOLOv3 network

Once the total vehicles count is obtained from the video frames at specified interval time, the traffic density value is to be communicated via internet. This information paves way to estimate the density of traffic in a specific time interval and the same is transferred to the central server system from the raspberry-pi.

B. Cloud Server

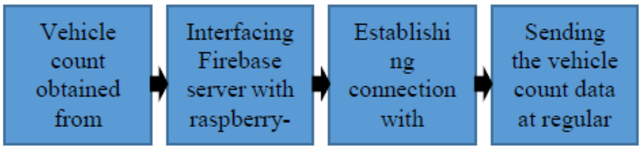

Figure 3. Flow diagram to establish connection with server and sending the vehicle count

The video camera communicates to the user cloud server system through the internet following the pattern as illustrated in Figure 3. First the obtained traffic density count from the YOLOv3 stored in a variable, which then updated at regular interval of time. Python programming is used to interface the firebase server with raspberry-pi. The processed information has to be sent to traffic office control room through the internet which demands internet connectivity for the raspberry-pi.

Python programming enables the communication with the controller while using firebase real time database. Once the interfacing has been done, the connection needs to be established with the database on firebase server. The database is connected through the Application Programming Interface (API) by sending the credential keys at the back of URL. Finally, the vehicle density is sent to real time cloud server after establishing the connection to database. The count and the time stamp both are encoded in the JSON format and sent to the firebase cloud server. This process is repeated at regular time interval to monitor the real time traffic flow.

IV. EXPERIMENT RESULTS AND ANALYSIS

Two datasets were employed in training the YOLOv3. At the first place, BIT-Vehicle dataset containing 9580 vehicle images was used to train the model. It includes six vehicle types the number of each type is given respectively. Then the model is retrained on the Comp Cars dataset which was provided by Stanford University. The Comp cars contains 40,000 images which was collected from surveillance cameras including day and night scenes. The training and the validation datasets are the results of splitting the BIT vehicle dataset into two with the ratio of 8:2. This model is again re-oriented with comp cars through the split datasets as mentioned above.

From 50020 number of iterations, the initial learning rate is considered as 0.001, then the learning rate is split into 40,000 and 45000 times, the mini-batch is assumed to be 16, and the subdivision and weight coefficient is set to 4 and 0.0005 respectively and the momentum coefficient is set to 0.9. At the time of training, the input size of the model is altered, so that final model has effective detection outcome for the images of various sizes. The proposed algorithm is evaluated by making use of the test sets obtained from the Comp cars. The PR curve is drawn in accordance to the accuracy and recall. As shown in table 1, the mAP of the YOLOv3 model was the highest, reaching 95.60%. The average detection speed was 0.035s, which means that the model could deal with about 26 pictures in 1 second.

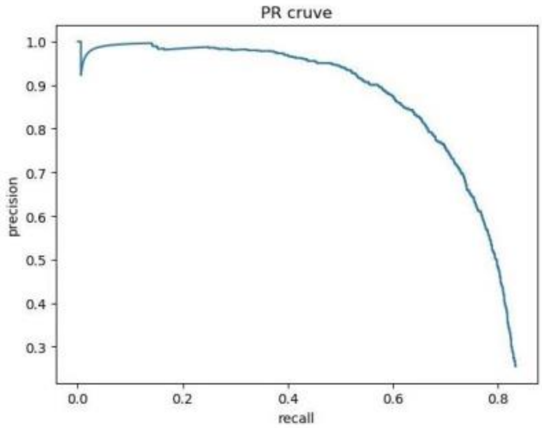

Figure 4. PR curve of improved YOLOv3

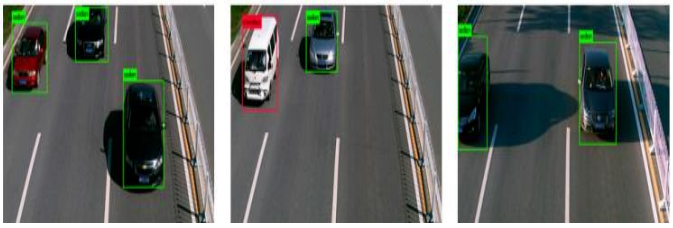

Figure 5. Detection results of improved YOLOv3 using BIT-Vehicle dataset

Table 1: Result using test data of Comp Cars

Model	mAP
Improved YOLOv3	95.6
YOLOv2	94.7
Faster R-CNN + ResNet	90.62

V. CONCLUSION

This research focusses on the improved YOLOv3 model was proposed for vehicle detection. For obtaining the optimal anchor boxes, the vehicle bounding boxes on the training dataset were clustered with k-mean clustering. Next the loss function is computed to decrease the different scaling of objects in the images. Then, for attaining the features, the network was designed to remove the high repeated convolution layers. Based on the experiments, the results showed that, mAP of improved YOLOv3 could reach 95.6%. The proposed methodology also showed good generalization ability using a dataset different from the training dataset. Therefore, the proposed network is effective for vehicle detection which could detect the vehicle fast and accurate, this helps the IoT devise to process the traffic density analysis faster.

REFERENCES

1. Cao, X.; Wu, C.; Yan, P.; Li, X. Linear SVM classification using boosting HOG features for vehicle detection in low-altitude airborne videos. In Proceedings of the 2011 IEEE International Conference Image Processing (ICIP), Brussels, Belgium, 11–14 September 2011; pp. 2421–2424.
2. Pan, C.; Sun, M.; Yan, Z. The Study on Vehicle Detection Based on DPM in Traffic Scenes. In Proceedings of the International Conference on Frontier Computing, Tokyo, Japan, 13–15 July 2016; pp. 19–27.
3. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv* **2016**, arXiv:1506.01497v3. [CrossRef] [PubMed]
4. Cai, Z.; Fan, Q.; Feris, R.S.; Vasconcelos, N. A unified multi-scale deep convolutional neural network for fast object detection. In Proceedings of the 2016 European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 354–370.
5. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
6. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525.

7. T. e. a. Osman,"Intelligent traffic management system for a cross section of roads using computer vision," in Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual, 2017.
8. Mahesh Lakshminarasimhan, "IoT Based Traffic Management System", Mach 2016.
9. H. O. Al-Sakran, "Intelligent traffic information system based on the integration of Internet of Things and Agent technology," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 6, no. 2, pp. 37–43, 2015.
10. S.Rane, A. Dubey and T. Parida, "Design of IoT Based Intelligent Parking System Using Image Processing Algorithms," IEEE International Conference on Computing Methodologies and Communication, pp. 1049–1053, 2017.
11. S. Nagmode and S. M. Rajbhog, "An Intelligent Framework for Vehicle Traffic Monitoring System using IoT," IEEE International Conference on Intelligent Computing and Control, 23-24 June 2017.
12. S. V. Bhate, P. V. Kulkarni, S. D. Lagad, M. D. Shinde and S. Patil, "IoT based Intelligent Traffic Signal System for
13. Emergency vehicles," International Conference on Inventive Communication and Computational Technologies, pp. 788–793, 2018.

AUTHOR'S PROFILE

S. Saravanan, received B.Sc. degree in Computer Science as well as M.Sc. degree from Bharathidasan University and M.Phil. degree from Annamalai University, India. He joined Annamalai University in the year 2001. At present he is with the Department of Computer and Information Science, Annamalai University. He has almost 19 years of teaching experience. He is currently pursuing Ph.D. degree in the Department of Computer and Information Science, Annamalai University. His research interests include Image processing, Machine Learning and Deep Learning.

Dr. K. Venkatachalapathy, Having more than 27 years of Teaching and 2 years Industry experience. He has more than 113 research papers in National/International Journals/Conference. He has guided 12 M.Phil. scholars and more than 10 Ph.D. Research scholars in Digital Image Processing, Computer Networks, Data Mining.

Retrieval Number: B4255129219/2019©BEIESP
DOI: 10.35940/ijeat.B4255.129219
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication