A Survey of approaches for vehicle traffic analysis

Rajesh Babu C[1], Anirudh. G[2], Naveen Venkatesh K[3]

[1]*Assitant Professor, Dept of Computer Science, SRM Institute of Science and Technology, Kattankulatur, Tamil Nadu, India*

[2,3]*Student, Dept of Computer Science, SRM Institute of Science and Technology, Kattankulatur, Tamil Nadu, India*

---***---

Abstract - *The ever-increasing traffic congestion especially in the urban area calls for better traffic surveillance and analysis sytems to be deployed. Vehicle traffic data accumulated over a period of time can be used to find traffic patterns and gain useful insights which can be used for improving the traffic management. Existing techniques for traffic analysis include magnet based loop detectors embedded inside the road, infra-red sensors on the side provide limited vehicle counting and traffic surveillance. Also, the sensors are prone to physical damage over a period of time, which reduces their functionality and accuracy. Several techniques perform a reliable job during the daylight conditions but falls short in the low light due to illumination change which drastically reduces the accuracy. The vision and video based deep learning techniques use object segmentation and several holistic approaches to arrive at the objective. Empirical evaluation on public-scale confirm that these methods can achieve a high real-time performance and accuracy while providing insightful data for further analysis. We analyze the most commonly used algorithms for vehicle traffic analysis*

Key Words: **Traffic analysis, Convolutional neural network, deep learning, yolo, video based system**

1.INTRODUCTION

The rationale behind this study was to understand the existing ways that are used to analyze vehicle traffic and to weigh up the pros and cons of the methods studied.

The existing methods have their own advantages which can be quite helpful but also has some disadvantages which can undermine the purpose of the implementation or may have a specific limitation under certain conditions.

While the existing methods may provide good real-time results, there is a trade-off made between the speed of determining the results and the accuracy in which the result is determined.

These trade-offs have significant impact on the overall performance of the system and cause significant delay or it provides inaccurate results [1]. Such inefficient systems when used in high density traffic areas drift off from their main purpose of installation, which is to monitor and analyze the traffic in real-time in a given area for a period of time. This is because, in high traffic region, the vehicle density for a given area is too high for the system to analyze in real-time and as a result produces inaccurate results.

While, in low traffic areas where the vehicle speed is very high, such as highways, the movement of vehicles are so fast for the camera to capture and as a result produces incomplete results. [2] The proposed system aims to provide a balanced working in all kinds of traffic conditions while minimizing the trade-offs. The major existing systems include Magnetic Loops, Microwave RADAR, Laser Based Systems, Infrared Devices, Ultrasonic devices, Anisotropic Magneto-Resistive (AMR) Magnetic Sensors.

Various image based methods have also been implemented, some of them include : Using Edge Detection, Blob Tracker Detection, Background Subtraction, Expectation Maximization Algorithm. Many of these methods have been employed in the past and have been successful to an extent in determining the vehicle traffic density[2].

Magnetic Loops, Microwave RADAR, Laser Based Systems, Infrared Devices, Ultrasonic devices, Anisotropic Magneto-Resistive (AMR) Magnetic Sensors are hardware based methods were the sensors are embedded in the road, which are susceptible to physical damage.

Edge Detection, Blob Tracker Detection, Background Subtraction, Expectation Maximization Algorithm are software based methods which uses the video feed from the surveillance camera to analyze the vehicle density at any given point of time

These methods suffered because of various trade-offs which made the final result either incomplete or inaccurate.

The convolutional neural networks based system, known as yolo (you only look once) is considered. In this system, a neural network is pre-trained to detect vehicles and pedestrians.

Video produced by a surveillance camera would be fed as an input to the model [3]. It will then analyze the video frame by frame processing each frame exactly once. The model would then create bounding boxes to identify the vehicle and count the number of vehicles in the frame.

The predicted result of the entire project is to reduce the time required to obtain the result with minimal loss of accuracy and speed. We use YOLO, because of its ability to detect stationary vehicles and to eliminate the problems caused by illumination changes.

The main advantage of YOLO is its flexibility in providing the user with the choice of speed-accuracy trade off which can be balanced out by arriving at a sweet spot by using trial and error method.

This achieved by using a single convolutional neural network that predicts multiple bounding boxes simultaneously and calculates the class probabilities for those boxes.

These boxes then detect the objects which is used to analyze the vehicle density in a given area for a given period of time.

2. VEHICLE TRAFFIC ANALYSIS TECHNIQUES

2.1 Edge Detection

An image's edge is the boundary of the image where occurs changes in various parameters such as reflection of surface, changes in illumination and the distance between the viewer and the visible surface.

Variations in physical aspects of the image can occur in various ways including the changes in parameters such as color, intensity and texture of the image[4].

The edge of an image can contain vast variations. They provide important visual information as they are caused by major changes in photometrics, physical or variations in geometry of the scene object [5].

Physical edges are caused by changes in the surface reflectance, illumination of light, physical orientation and depth of the scene.

Fig -1: Edge Detection of a vehicle in the road

The most important aspect of is to detect and count vehicles in a daylight environment by using real time traffic density changes by using various differential techniques and using lane dividing methods to effectively identify and count vehicles[6]. The result of vehicle counting system can be used to track vehicles in high occlusion areas. Vehicle density at different illuminations can be found to an extent.

The advantages are that it is one of the simplest and efficient methods available for day time vehicle density detection. Noises can be removed, so the image is directly detected and noises are filtered out without pre-processing. There are no complex mathematical operations, hence high speed performance.

The Disadvantages are that there is no mechanism for shadow removal and handling high occlusion areas. It produces comparatively poor results. There is chance for multiple vehicles to be counted as single object in high occlusion areas

2.2 Background Subtraction

The method of extracting moving foreground objects from a static background is called as background subtraction.

Any foreground object can be detected by producing difference of the sequence of background model. The result from this subtraction can be further used for tracking specific targets and for vehicle movement detection.

Fig -2: Original and Background subtracted image

Its main use cases involve vehicle detection and monitoring, recognizing human actions, tracking human computer interaction, tracking movement of various objects and digital forensics[7].

Advantages include that is algorithm is extremely easy to implement, ease-of-use and its very fast.

Disadvantages include that the framing accuracy depends on object tracking speed and frame rate of the video. It also has relatively very high memory requirements.

The model does not give good results in the following conditions – if a bi-modal background exists, if there are many slow moving or static objects in the frame, If the video frame rate is low but movement of objects are fast and if there is abrupt changes lighting condition of the scene from time to time [8]

2.3 Blob Detection

The blob can be defined as a square either fully or partially filled and any square that can be reached from the original square by either horizontal or vertical movements. By traversing through pixels of the image, a blob can be detected.

Blob filtering is carried out to filter out the static blobs which form the background of an image[9].

Two types of blob detection techniques can be used –

(i) differential method which is based on the derivative function with respect to the image position

(ii) local extrema methods which is based on finding the local minima and maxima of a function.

Using either one of the two blob detection methods, a blob list is created to find out which blobs contain only the vehicles.

Advantages include that it can be used to track multiple vehicles at the same time in one iteration. Removes duplicate vehicles and shadows[10].

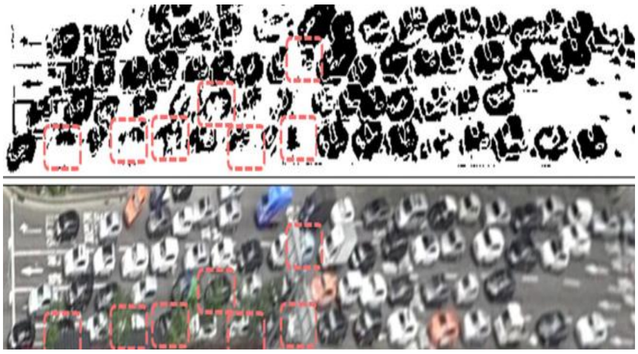

Fig -3: Blob tracker implementation

Disadvantages are that the segmentation quality of moving vehicles is poor in a cluttered environment. Cannot detect stationary vehicles.

2.4 YOU ONLY LOOK ONCE

YOLO is a convolutional neural network for object identification and detection. The object detection task is a two-step process which consists of (i) determining the object location on the image and (ii) classifying those objects[1].

The existing methods such as R-CNN used an execution pipeline to perform this task that involves multiple steps to complete. Therefore, such an approach very slow to execute and as a result optimization becomes complicated as each individual component must be trained separately.

YOLO architecture is similar to that of a fully convolutional neural network and passes the image (n x n) once through the FCNN and the output obtained is (m x m) prediction [1]. YOLO employs a single convolutional network that has the ability to predict multiple number of bounding boxes simultaneously and then calculate the class probabilities for each box . The YOLO algorithm can train on full images and as a result it optimizes the detection performance.

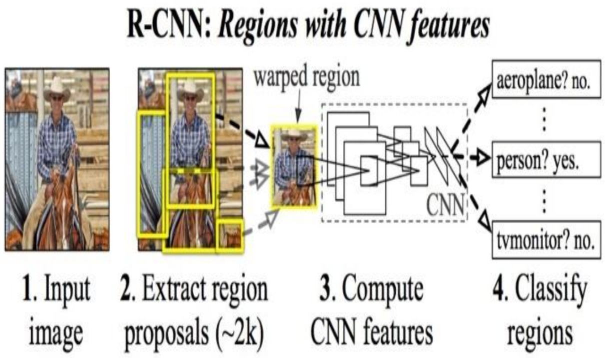

Fig -4: Regions with cnn features

YOLO offers multiple benefits over the existing methods of object and vehicle detection. Firstly YOLO is very fast . Secondly YOLO takes the entire image into consideration while making the predictions. Unlike the existing techniques such as sliding window and region proposal, YOLO takes the full image for training and test and so it passively encodes context-based information on the classes and their appearance.

Each and every bounding box is predicted by using features from the entire image. It can also simultaneously predict the bounding boxes for all classes. This implies that the neural network can analyze all of the objects in an image when using a full image[7].

The YOLO's design allows to maintain high precision while enabling real time end-to-end training.

The network divides the input image into an N x N matrix-grid. If the center of an object comes under a grid cell, that particular grid cell is responsible for detection of that object.

Each and every grid cell produces the bounding boxes and generates the confidence scores for them. These confidence scores reflects the confidence level of the model as the box contains an object. The confident score of the cell should be zero, if no object exists in that cell [5].

In this model we will be giving a video as the input. The model works on the video frame by frame and analyses each frame exactly once and detects the objects/vehicles present in that frame. It Takes a count of the number of vehicles and stores the count.

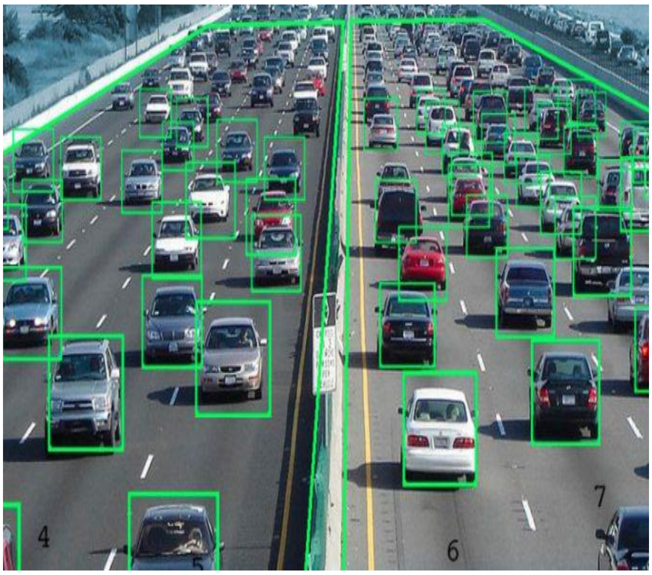

Fig -5: Working of Yolo

This vehicle count is used to calculate the vehicle density at a particular place [4].

Since YOLO works in real time, the count produced by the model can be reliably used to analyze the real time traffic scenario in that particular area. The tradeoff between speed and accuracy and be achieved by changing the size of the model. No re-training of the model is required. Because of this, cost of re-training is completely removed which thereby reduces the overall cost of operation of this model.

Advantages of YOLO are that It can detect stationary vehicles as well as moving vehicles, It is very fast and accurate and Works in real-time.

Disadvantage is that that initial training requires high computational power.

3. CONCLUSION

This paper analyses the major vehicle detection algorithms. As discussed in the paper, each algorithms have their pros and cons. We have studied the above mentioned algorithms in detail and can conclude that the yolo algorithm offer major benefits over the rest of the algorithms.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition, 2016, pp. 779–788

[2] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.

[3] C. Ozkurt and F. Camci, "Automatic traffic density estimation and vehicle classification for traffic surveillance systems using neural networks,"*Math. Comput. Appl.*, vol. 14, no. 3, pp. 187–196, Dec. 2009

[4] C. Ozkurt and F. Camci, "Automatic traffic density estimation and vehicle classification for traffic surveillance systems using neural networks,"*Math. Comput. Appl.*, vol. 14, no. 3, pp. 187–196, Dec. 2009

[5] Vehicle counting for traffic management system using YOLO and correlation filter, Asha C S, A V Naarasimhan, 2018 IEEE Conference

[6] P. Viola and M. J. Jones. Robust real-time face detection. International journal of computer vision, 57(2):137–154,2004

[7] M. Sun, Y. Wang, T. Li, J. Lv, J. Wu, Vehicle Counting in Crowded Scenes with Multichannel and Multi-task Convolutional Neural Networks, *J. Vis. Commun. Image R*

[8] A Computer Vision Based Vehicle Detection and Counting System, Nilakorn Seenouvong , Ukrit Watchareeruetai and Chaiwat Nuthong, 2016 IEEE Conference

[9] B. Hariharan, P. Arbel´aez, R. Girshick, and J. Malik. Simultaneous detection and segmentation. In Computer Vision–ECCV 2014, pages 297–312. Springer, 2014

[10] An obstacle avoidance algorithm with spatial and temporal constraints for visually impaired, Amutha, B., Nanmaran, K, International journal of apllied research engineering, 2015