International Journal of Innovative Research in Engineering & Management (IJIREM)
ISSN: 2350-0557, Volume-9, Issue-2, April 2022
https://doi.org/10.55524/ijirem.2022.9.2.106
Article ID IJIRP-2305, Pages 661-666
www.ijirem.org

Vehicle Speed Estimation and Traffic Tracking System Using Machine Learning

R. Sai Chandu[1], G. Venkata Krishna[2], T. Karthik[3], Sk. Firowz[4], Mr. Jilani Basha[5], and Dr. K. G. S Venkatesan[6]

[1,2,3,4,5,6] Department of Computer Science and Engineering, PACE Institute of Technology& Sciences, Vallur, Ongole, Andhra Pradesh, India

Correspondence should be addressed to R. Sai Chandu; 18kq1a05i0@pace.ac.in

ABSTRACT- In this work, we deploy a real-time method for classifying vehicles and estimating their speeds using footage captured by traffic cameras along motorways. Basic techniques in traffic analysis include the forecasting of traffic flows, the discovery of anomalies, the re-identification of vehicles, and the tracking of moving vehicles. One of the most actively studied areas of these applications is traffic flow prediction, often known as vehicle speed estimation. In this work, we estimate vehicle speeds within classes using feature tracking and neighbor discovery techniques.

KEYWORDS- Speed Estimation, Traffic Camera, Feature Tracking, Vehicle Classification

I. INTRODUCTION

The primary objectives of Intelligent Transportation Systems (ITS) include the reduction of power consumption and emissions, the elimination of traffic delays and congestion, and the improvement of overall road safety. Information provided by ITS helps daily users as well as transportation organizations better their monitoring, management, and control of traffic flow, which in turn benefits the users. As a component of an ITS, determining a vehicle's geo-referenced position and speed using computer vision technologies is the purpose of the work that is being presented here. In our earlier work [1], in which we constructed an algorithm for describing traffic flow but did not include any classification capabilities, we recognized some of the benefits of gravitating toward a Computer Vision approach. These benefits were recognized after we had already developed an algorithm for describing traffic flow. In this particular piece of writing, our focus is on the classification problem. In the following section (section 2), we will discuss the Computer Vision methods and Machine Learning approach that were utilized for the work. We provide the measurement findings that were attained by applying these strategies to a video that was acquired from a highway traffic camera. In the final, section 3, we talk about the outcomes and any changes that could be made in the future.

In this study, we offer a strategy for solving Track1, which may be found here.

The detection and tracking of vehicles are extremely important to our model. Since there are no labeled data provided for this challenge, it is difficult to train a vehicle detection model using grazing data. Instead, we make use of transfer learning and carry out inferences on our dataset by employing the 3D Deformable model. This allows us to locate vehicles. We thought about using the model [2]. instead of the 3D Deformable model because it had a comparable presentation in the 2017 competition. To be able to evaluate the performance of these two models in relation to the target reference data, we first extract all of the frames from one of the Track 1 movies and then test the models' performance on the frames. Specifically, we compare the models' ability to detect vehicles based on their mean Average Precision (mAP) scores. According to the findings of the research, the 3D Deformable model gets 74% mAP [3], which is a score that is greater than the model developed [4].

Our tracker uses a mechanism called detect-then-track for its methodology. The effectiveness of the tracker is extremely predicated on the degree of precision achieved by the detection. In each frame, we pull out peculiar characteristics from the automobiles that have been spotted, and then in the following frame, we identify those vehicles. The information required for calculating an accurate estimate of the vehicle's speed can be given by the movement of these components inside the frame.

II. MATERIALS AND METHODS

Intelligent Transportation Systems (ITS) have several basic goals, the most important of which are the decrease of overall power consumption as well as emissions, the elimination of traffic delays and congestion, and the improvement of general road safety [5]. The information that is produced by ITS assists daily users as well as transportation companies in improving their monitoring, management, and control of traffic flow. This, in turn, is beneficial for the daily users. The goal of the work that is being presented here is to determine a vehicle's geo-referenced position and speed using computer vision technology. This is intended to be a component of an ITS (Intelligent Transportation System). In our prior work [6], in which we created an algorithm for characterizing traffic flow but did not include any classification capabilities, we recognized some of the benefits of migrating toward a computer vision approach.

These benefits include: recognizing some of the benefits of gravitating toward a computer vision approach.

recognizing some of the benefits of gravitating toward a computer vision approach. After we had already built an algorithm for characterizing the flow of traffic, we didn't realize these benefits until much later. In this particular piece of writing, the categorization problem will serve as the primary focus of attention. In the section that comes after this one, we will talk about the Computer Vision algorithms and the approach to Machine Learning that were applied for the job. We provide the measurement findings that were achieved by applying these methodologies to a video that was obtained from a highway traffic camera. The footage was captured by the camera as it was moving down the highway [7].

Within the scope of this investigation, we present a method for resolving Track1, which may be located here. The identification and monitoring of moving vehicles are of the utmost significance to our concept. It is difficult to train a vehicle detection model utilizing grazing data because there are no labeled data available for this challenge. Instead, we draw conclusions from our dataset by utilizing transfer learning and the 3D Deformable model. This enables us to pinpoint the location of automobiles. Because of its comparable appearance in the 2017 competition, the model that was developed by Bhandaryet al. came into consideration as a potential replacement for the 3D Deformable model. In order to evaluate the performance of these two models in regard to the target reference data, we must first extract all of the frames from one of the Track 1 videos and then test the models' performance on the extracted frames. This will allow us to analyze the models' performance in relation to the target reference data. To be more specific, we evaluate the models' capacity to recognize automobiles by contrasting the mean Average Precision (mAP) values that each receives [8]. The findings of the study indicate that the 3D Deformable model receives 74% mAP, which is a score that is higher than the model that was built by Bhandary et al.

The methodology behind our tracker is based on something called a detect-then-track mechanism. The degree of precision that can be attained in the detection is an extremely important factor in determining how useful the tracker will be. At the end of each frame, we go on to the next one and identify the vehicles based on the distinctive traits that we extracted from the previously spotted automobiles in the previous frame. The movement of these components within the frame can provide the information that is necessary for computing an accurate estimate of the speed at which the vehicle is traveling.

Step one is to gather videos of traffic from various web cameras.

As was detailed in Section 2.1, the purpose of this investigation was to collect traffic videos from online cameras in order to use them for image data training and vehicle classification verification.

In this second part, you will count vehicles by employing GMM and a virtual detection zone.

Object detection and recognition are carried out in order to get accurate counting of vehicles in real time. For the purposes of carrying out vehicle counting and speed estimate, respectively, a virtual detection zone and a virtual detection lane line are utilized, as detailed in Sections 2.2 and 2.4, respectively.

In this third and last step, you will use the YOLOv3 and YOLOv4 algorithms to perform vehicle categorization and speed estimation.

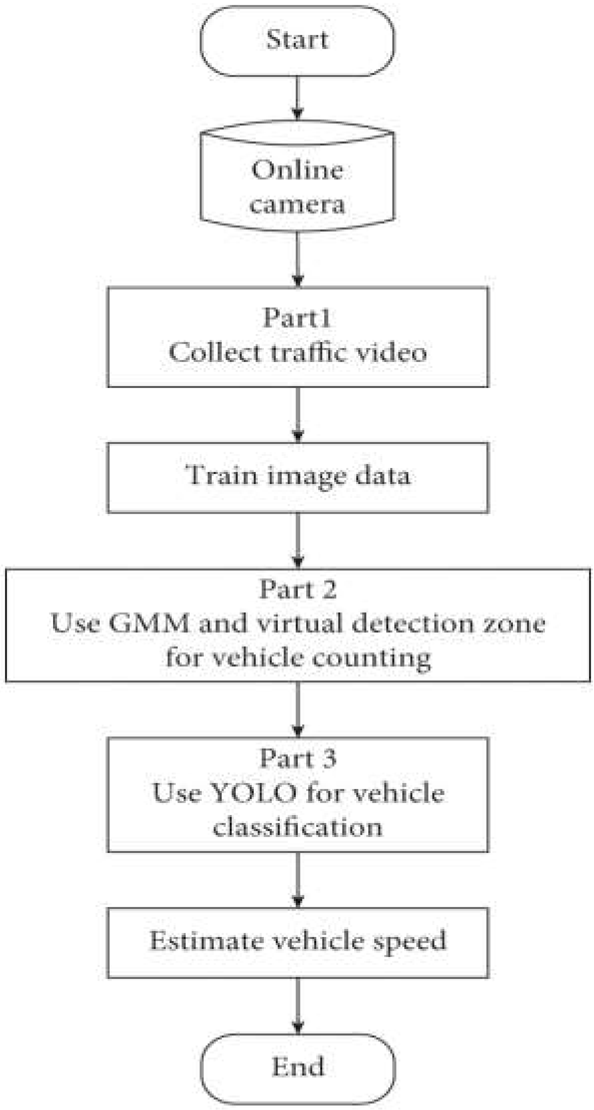


Figure 1 : Flowchart of the vehicle counting and classification process

A. Data Set Preparation

The data set that was used in this investigation was compiled by gathering recordings of traffic that were collected with internet cameras that were installed along a large number of roadways in Taiwan. A script was used to extract image data from the traffic movies, and an open-source software tool named "labelling" was used to label the images once they were extracted. This study splits training into sedans, trucks, scooters, buses, hlink vehicles, and flink automobiles. Table 1 shows these six vehicle types' lengths. This study splits the training procedure into six dimensions based on Taiwan's Directorate General of Highways, Ministry of Transportation and Communications (MOTC)'s common road vehicles. This study used YOLO without considering vehicle length [8].

Table 1: Vehicle classification

Class	Vehicle	Length (m)	Image
1	Sedan	3.6–5.5	
2	Truck	>5.5–11	
3	Scooter	1–2.5	
4	Bus	7–12.2	
5	Hlinkcar	15–18	
6	Flinkcar	18–20	

B. Vehicle Counting

In order to get an accurate count of the cars in a complicated environment, a GMM is employed for the contextual subtraction, which pinpoints the areas that contain moving items. As a result of the GMM's high level of dependability in the backdrop concept and the foreground split-up process, the defining attributes of a moving object can be more readily identified in video surveillance. Each video has a predetermined virtual detection zone that is then utilized for the estimation of vehicle characteristics. When the vehicle enters a virtual detection zone and crosses a virtual detection lane line, the GMM is used to count the number of vehicles in the area [9]. Figure 2 provides an explanation of the counting window for vehicles.

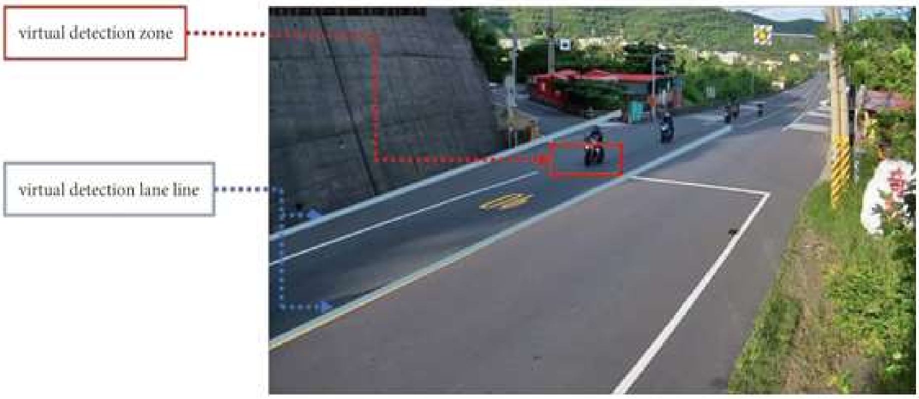
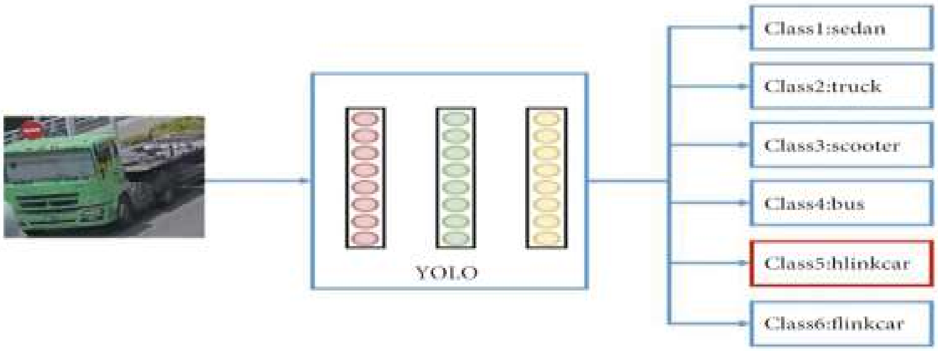


Figure 2: Object detection window

C. Classifying Vehicles

This technique employs the YOLO classification algorithm in order to divide automobiles into six distinct categories. The authentication method is utilized in the process of confirming the vehicle organization in the videos that have been collected. When checking the capabilities of the vehicle's organization, a visual classifier that is based on the YOLO method is utilized. Figure 3 depicts the architecture of the visual classifier that is based on the YOLO algorithm and is used for classifying each vehicle into one of six categories. These categories are shown in the figure. During the process of preparation, when a vehicle that fits into one of the six classes is found, all bounding boxes are extracted, their classes are manually labeled, and the data from the manually labeled bounding boxes are supplied to the YOLO model so that the vehicle can be classified using those bounding boxes [10]



.

Figure 3: Architecture of visual classifier based on the YOLO algorithm for verifying the vehicle classification.

D. Calculation of the Speed

In addition, the real-time speed of the vehicle is built for this investigation. Figure depicts the video pictures that were occupied along the route in a manner that is parallel to the length of the automobile (represented by the y-axis) and comparable to the size of the car (represented by the x-axis). First, according to the scale in the video, the length of the yellow line that is specified as "L" in the red circle is 4 meters, which is in accordance with the regulations regarding traffic. On the route that is going to be tested, a virtual finding zone, also known as a blue box, can be drawn with the help of a GMM. The car's chassis, denoted by the letter C, is denoted by the letter Ct, and the green box serves as the chassis.

$$u_0 = \frac{\overline{L_{AB}}(px)}{\overline{L_{AB}}(m)},$$

$$\alpha = \frac{u_0}{u_0'},$$

$$\text{If}\,\alpha > 1, \quad u_i = u_0\alpha^i, u_j = \frac{u_0}{\alpha^j},$$

where u_0 is the scale, i is the scale of the blue box, j is the scale of the green box, px is the length of the video, and m is the actual length. The parameter α denotes the increase or decrease in relationship of the scale per unit length on the y-axis. If $\alpha > 1$, the speed calculation is performed using equation

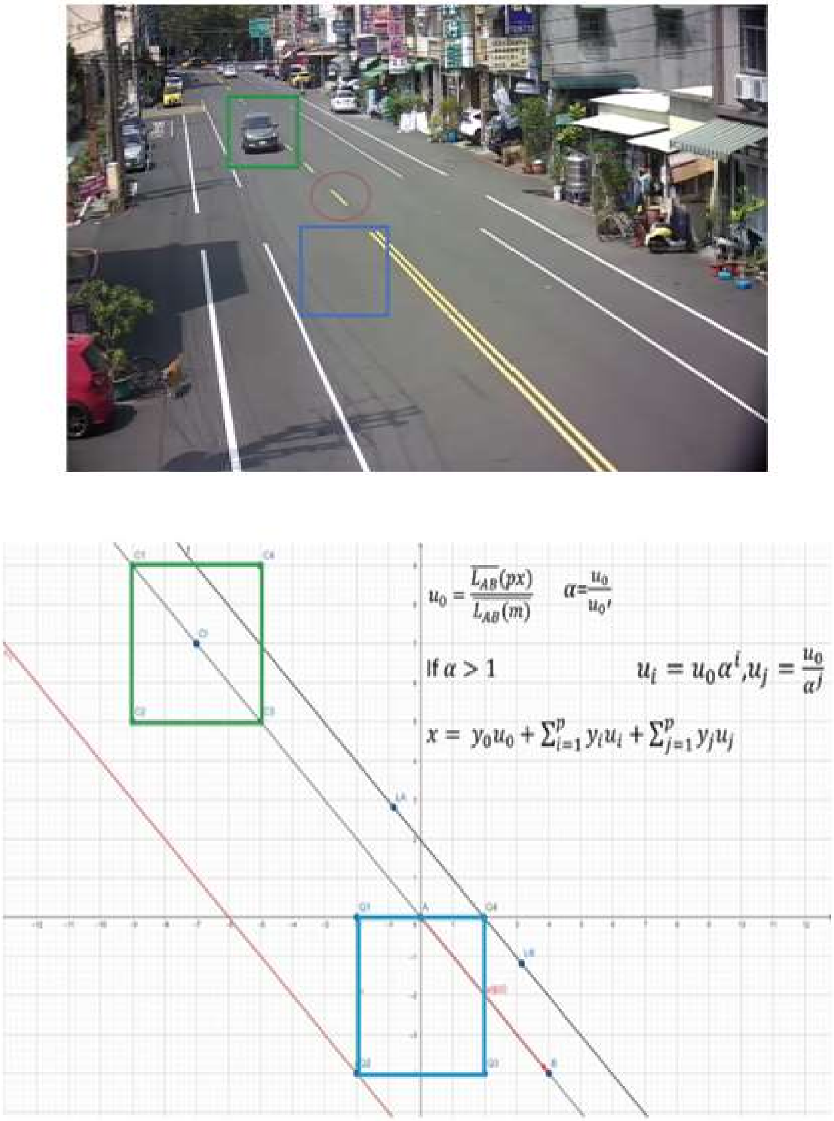


Figure 6: Speed estimation diagram ((a): video image; (b): control scale).

To calculate the parallel L line segment of Ct (referred to as L^*), the algorithm computes the L^* distance y between A and B. Then, it restores y from its scale relationship with the actual line segment length x, where x denotes the distance traveled by the vehicle in Q.

$$x = y_0u_0 + \sum_{i=1}^{p} y_iu_i \sum_{j=1}^{p} y_ju_j.$$

In the calculation process, the program is used to determine the frame rate of the video and calculate the number of frames for which the vehicle travels in Q (referred to as p). Equation is used to find the travel time of the vehicle from A to B in Q.

$$t = \frac{p}{\text{fps}},$$

$$v = \frac{x}{t},$$

$$v' = x \times \frac{3.6}{t}.$$

Equation is used for calculating vehicle speed. After unit conversion (m/s to km/h), Equation provides the vehicle speed.

III. MAVD AND GRAM-RTM DATA COMPARISON RESULTS

The MAVD traffic records set and the GARM Road-Traffic Monitoring (GRAM-RTM) dataset were utilized in order to evaluate the effectiveness of the proposed method regarding the counting of vehicles. The films were captured using a GoPro Hero3 camera at a frame rate of 30 frames per second and a resolution of 1920 pixels by 1080 pixels. After reviewing ten films, we found that the proposed method has a 93.84% accuracy rate for counting vehicles at ten in the morning for the MAVD traffic data set. Table 2 presents the results of the projected method's application of the MAVD traffic data set to the classification of vehicles.

Table 2: Vehicle classification result of the proposed method using MAVD traffic data set

Video no.	Total number of vehicles			Number of counted vehicles		
	S	*L*	Total	*S*	*L*	Total
1	8	0	8	8	0	8
2	5	0	5	4	0	4
3	4	2	6	3	2	5
4	4	0	4	3	0	3
5	3	0	3	3	1	4
6	1	0	1	1	0	1
7	7	2	9	7	2	9
8	11	0	11	10	0	10
9	9	0	9	9	0	9
10	9	0	9	8	0	8

IV. CONCLUSIONS

The MAVD traffic records set and the GARM Road-Traffic Monitoring (GRAM-RTM) dataset were used in order to determine whether or not the proposed approach for counting vehicles was an accurate representation of actual traffic patterns. These movies were recorded with a GoPro Hero3 camera at a frame rate of 30 frames per second and a resolution of 1920 pixels by 1080 pixels. After going through ten different films, we discovered that the method that was suggested has an accuracy rate of 93.84% when it comes to counting vehicles at ten in the morning for the MAVD traffic data set. The results of applying the proposed method to the MAVD traffic data set in order to classify cars are presented in Table 2, which can be found below.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Y. Mo, G. Han, H. Zhang, X. Xu, and W. Qu, "Highlight-assisted nighttime vehicle detection using a multi-level fusion network and label hierarchy," Neurocomputing, vol. 355, pp. 13–23, 2019.

[2] D. Feng, C. Haase-Schuetz, L. Rosenbaum et al., "Deep multi-modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 3, p. 3, 2019.

[3] Z. Liu, Y. Cai, H. Wang et al., "Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions," IEEE Transactions on Intelligent Transportation Systems, 2021.

[4] Y. Qian, J. M. Dolan, and M. Yang, "DLT-NET: joint detection of drivable areas, lane lines, and traffic objects," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 11, pp. 4670–4679, 2020.

[5] Y. Cai, L. Dai, H. Wang et al., "Pedestrian motion trajectory prediction in intelligent driving from far shot first-person perspective video," IEEE Transactions on Intelligent Transportation Systems, pp. 1–16, 2021.

[6] Á. Llamazares, E. J. Molinos, and M. Ocaña, "Detection and tracking of moving obstacles (DATMO): a review," Robotica, vol. 38, no. 5, pp. 761–774, 2020.

[7] C. Liu, D. Q. Huynh, Y. Sun, M. Reynolds, and S. Atkinson, "A vision-based pipeline for vehicle counting, speed estimation, and classification," IEEE Transactions on Intelligent Transportation Systems, pp. 1–14, 2020.

[8] Y.-Q. Huang, J.-C. Zheng, S.-D. Sun, C.-F. Yang, and J. Liu, "Optimized YOLOv3 algorithm and its application in traffic flow detections," Applied Sciences, vol. 10, Article ID 3079, 2020.

[9] W. Liu, D. Anguelov, D. Erhan et al., "SSD: single shot multibox detector," in Proceedings of the Computer Vision—ECCV 2016, pp. 21–37, Amsterdam, Netherlands, October 2016.

[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.