

A VIDEO BASED VEHICLE DETECTION, COUNTING AND CLASSIFICATION SYSTEM

**Mahesh Suryawanshi*1, Akash Ghadage*2, Shubham Ingwale*3,
Prathamesh Sutar*4, Shubham Mudshingkar*5, Prof. R.S. Nejkar*6**

*1,2,3,4,5 Students, Department Of Computer Science & Engineering, Sanjeevan Engineering
And Technology Institute Panhala, Maharashtra, India.

*6 Assistant Professor, Department Of Computer Science & Engineering, Sanjeevan Engineering
And Technology Institute Panhala, Maharashtra, India.

ABSTRACT

Traffic monitoring on road needs to be done, the counting of the number of vehicles passing the road is necessary. It is more emphasized for highway transportation management in order to prevent efforts. Therefore, it is necessary to develop a system that is able to counting the number of vehicles automatically. Video processing method is able to counting the number of vehicles automatically. This system includes processes of video acquisition, frame extraction, and image processing for each frame. This project has a system of vehicle counting on wide road. This system employs of background subtraction and morphology methods on gray scale images for vehicle counting. This presents a developed system which achieves vehicle detection by using background image subtraction algorithm.

Keywords: Vehicle Dataset, Image Segmentation, Vehicle Detection, Vehicle Counting.

I. INTRODUCTION

The increasing number of vehicles makes a lot of pressure on roads capacity and infrastructure. Accordingly, the traffic management is too difficult and the traffic hazards increases at a high rate, moreover it causes a huge loss of life and property due to the road accidents. Vehicle detection systems play an important role to reduce the harmful side effects of traffic, as a part of many traffic applications such as road traffic control, traffic response system, traffic signal controller, lane departure warning system, automatic vehicle accident detection and automatic traffic density estimation. Image processing methods are able to process video files to detect moving objects. The field of image processing, especially the method of object tracking is already widely used to detect and classify all sorts of moving objects. The advantages of this system are to calculate the number of vehicles automatically by using a camera device. With this system will get data in real time and can be done for monitoring the density of traffic. Moving vehicles detection is usually carried out by using background image subtraction techniques. The key behind this kind of techniques is to first build a background model from a sequence of images in order to find the moving vehicles from the difference between that background estimation and the current frame. The weakness of this system is its accuracy is very dependent on the illumination, with good illumination this system will have a very good accuracy. However, if the illumination is not good, as at night, the accuracy of this system will decrease significantly.

II. RELATED WORK

Various approaches were made to develop such systems that can detect, count and classify the vehicles and can be used for traffic surveillance in intelligent transportation systems. This section covers the discussion about such systems and the knowledge about the methods used to develop such systems. Tursun, M and Amrulla, G [4] proposed a video based real-time vehicle counting system using optimized virtual loop method. They used real time traffic surveillance cameras deployed over roads and compute how many vehicles pass the road. In this system counting is completed in three steps by tracking vehicle movements within a tracking zone called virtual loop. Another video-based vehicle counting system was proposed by Lei, M., et al. [5]. In this system surveillance cameras were used and mounted at relatively high place to acquire the traffic video stream, the Adaptive background estimation and the Gaussian shadow elimination are the two main methods that were used in this system. The accuracy rate of the system depends on the visual angle and ability to remove shadows and ghost effects. The system's incompetency to classify vehicle type is the core limitation of the system. Bas et

al. proposed a video analysis method to count vehicles [10] based on an adaptive bounding box size to detect and track vehicles in accordance with estimated distance from the camera. The Region of Interest (ROI) is identified by defining a boundary for each outbound and inbound in the image. Although the algorithm is improved to deal with some weather conditions it is unable to track vehicles when they change their directions.

Mithun, N.C., et al proposed a vehicle detection and classification system using time spatial image and multiple virtual detection line [6]. A two-step K nearest neighborhood (KNN) algorithm is adopted to classify vehicles via shape invariant and texture-based features. Experiments confirm the better accuracy and low error rate of proposed method over existing methods since it also considers the various illumination conditions.

Habibu Rabi proposed a vehicle detection and classification for cluttered urban intersection [11]. In this system background subtraction and Kalman filter algorithm are used to detect and track the vehicles and Linear Discriminant Analysis classifier is used for proper classification of vehicles.

Detection of vehicles in a video-based traffic surveillance system is first and very important phase as it greatly impacts the other algorithms such as tracking and classification of the vehicles hence an accurate detection and segmentation of the foreground moving object is very important. Many of the techniques are used for foreground detection like frame differencing [12]. Frame differencing can be considered as the simplest foreground detection and segmentation method as it is based on the close relationship among the sequence of motion images.

A. Suryatali and V.B. Dharmadhikari in [25], proposed a Computer Vision based vehicle detection that counts and also classifies the vehicles in to heavy and light weight vehicles; object detection is accomplished by making use of Kalman filter for background subtraction and then OpenCV library is used finally to detect the object in processed frame. Nilakorn et al. [26], proposed vehicle detection and counting prototype which uses different steps for background subtraction and object detection then uses CV techniques such as thresholding, hole-filling and adaptive morphological dilation to remove noise and enhance the foreground objects in particular frame from video. Proposed system provides limited functionality for the objects appearing in detection zone if they are occluded or small.

III. PROPOSED METHODOLOGY

The system could be used for detection, recognition and tracking of the vehicles in the video frames and then classify the detected vehicles according to their size in three different classes. The proposed system is based on three modules which are background learning, foreground extraction and vehicle classification as shown in fig. 1. Background subtraction is a classical approach to obtain the foreground image or in other words to detect the moving objects.

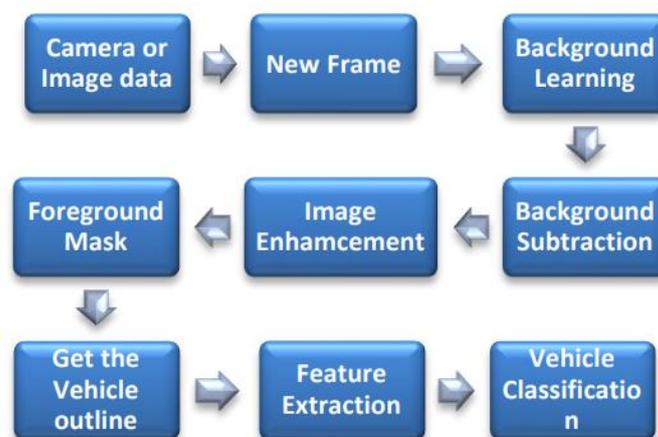


Fig.1. Block diagram of proposed vehicle detection, counting and classification system.

A. Background Learning Module

This is the first module in the system whose main purpose is to learn about the background in a sense that how it is different from the foreground. Furthermore, as proposed system works on a video feed, this module extracts the frames from it and learns about the background. In a traffic scene captured with a static camera installed on the road side, the moving objects can be considered as the foreground and static objects as the

background. Image processing algorithms are used to learn about the background using the above-mentioned technique.

B. Foreground Extraction Module

This module consists of three steps, background subtraction, image enhancement and foreground extraction. Background is subtracted so that foreground objects are visible. This is done usually by static pixels of static objects to binary 0. After background subtraction image enhancement techniques such as noise filtering, dilation and erosion are used to get proper contours of the foreground objects. The final result obtained from this module is the foreground.

C. Vehicle Classification Module

The third and the last module in the proposed system is classification. After applying foreground extraction module, proper contours are acquired. Features of these contours such as centroid, aspect ratio, area, size and solidity are extracted and are used for the classification of the vehicles.

IV. DETAILED METHODOLOGY

The first step of the proposed system is to grab a data video on which we want to perform the classification. After video selection, ROI is defined. ROI needs a careful human supervision because region of interest and imaginary line plays important role in classification. After ROI is defined, the system performs series of tasks i.e., applying background mask, subtracting mask, performing binary threshold, morphology using erosion and dilation, median blur, applying masked data to the frame, convert frame to gray scale. Contours are detected after these operations. Once contours are detected; system analyses the moments of the contours, marks the detected contours and centroid is calculated. If calculated centroid is in the range of the diagonal, system moves towards further operation for classification else system will be redirected towards the detection of contours again. The last step is the classification; the system classifies the vehicles with two different methods i.e., using SVM and with the CC.

The classification using SVM is used in which SIFT features are calculated for the contours and used as input to the SVM. Three types of vehicles are identified by SVM which are Low Transport Vehicle (LTV), Medium Transport Vehicle (MTV) and Heavy Transport Vehicle (HTV). SVM classifies the vehicle using the features extracted with the help of SIFT and then corresponding variables are incremented according to the output i.e., LTV, MTV and HTV. In addition, classification of the vehicle's using CC is done. Once the centroid calculated is in the range of diagonal; the properties of contours are extracted. The features extracted are compared with the assumed values and output is calculated. In the end the corresponding variables are incremented according to the output.

A. Region of interest

ROI is a particular portion of an image on which an operation is to be performed. ROI gives the flexibility to just work with in a particular region instead of manipulating the whole image. In proposed system, selection of region of interest is very important to reduce the false positives in the detection and classification of vehicles. Selection of ROI is pretty simple, once the video is started, the user has to press the "I" key on the keyboard to activate the input mode. Afterwards the user uses his mouse to select the four points on the video which defines the region of interest. Once selected, pressing of any key on the keyboard selects the region of interest, crops it and shows the new video feed on only that region. Fig. 2 shows ROI selection input mode. Notice the four green dots on the screen, these are the points defining the ROI and were placed using mouse clicks.



Fig.2. Region of interest selection input mode.

B. Background subtraction

Proper background subtraction is a vital pre-processing step in creation of any visual surveillance system as the accuracy of whole process of classification of the objects depends on it. The systems like visitor counter [27] in which a static camera captures the video of people entering a building and the system could count the number or a system where a camera captures the video of the vehicles on the road for the similar purpose.

The background subtraction could be an easy job if we already have an image of the background like the image of the building or the road. In cases defined above, background image could be removed and foreground objects could be obtained but most of the time the situation is varying. The backgrounds can be dynamic or initial information of the scene might not be available. Furthermore, the background subtraction becomes more difficult if the objects in the video also have shadows since they also move with the people or vehicles, then the normal background subtraction will detect the shadows as foreground objects too.

Several algorithms have been introduced for the situations like that; some of them are implemented in OpenCV such as Background Subtraction MOG [28] which use Gaussian distributions to create the model of the background in the image. It uses about 3 to 5 Gaussian distributions for this purpose. Another background subtraction implemented in OpenCV is called Background Subtractor GMG which is based on [27] and combines the background image estimation segmentation.

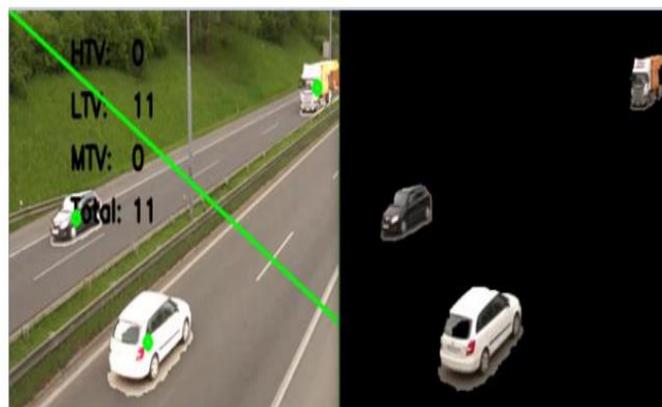


Fig.3. Frame of video before and after subtraction of background.

The algorithm used in the implementation of proposed system is called backgroundSubtractorMOG2. It is based on two studies [29] and [30] by Zikovic. One of the important features of this algorithm is that unlike [28] in which the number of distributions for the creation of background model are defined, Background Subtractor MOG2 uses an automated approach and selects an appropriate number of Gaussian mixtures for the pixel. Furthermore, algorithm is also better at handling the illumination changes in the scene. The algorithm also gives the ability to define if the shadow of the objects is to be detected or not. Note that the default settings of the implementation are set to detect the shadows.

C. Contour extraction

Contours are the boundaries of the shape which are used for the shape detection and recognition. The accuracy of the process of finding the contours can be defined as the canny edge detection performed on a binary image. OpenCV provide cv2.findContours() method to find the contours.

D. Counting vehicles

Vehicle counting can be done by checking if the centroid of a vehicle has touched or crossed the imaginary line in ROI. Imaginary line is the line that diagonally appears by connecting two ROI points. Once the centroid of a vehicle in ROI crosses the imaginary line, the system counts the vehicle. From fig. 4, it is noticeable that when centroid of vehicle touches the imaginary line in ROI, it increments in the variable "Total" and variable of respective category.

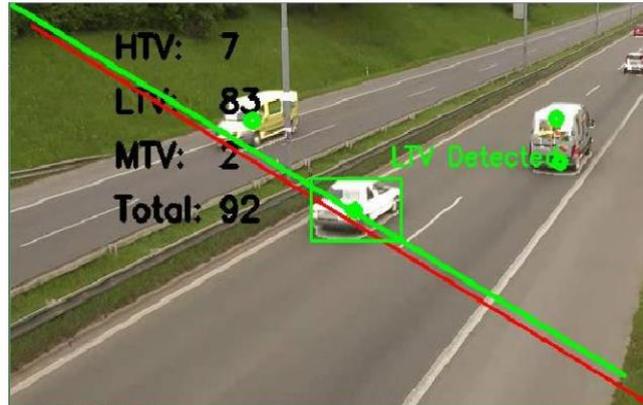


Fig.4. Increment in corresponding variables when vehicle centroid touches imaginary line in ROI (Mask Applied)

E. Classification

The classification is done using two different methods in implementation of the system which are: i) Classification using CC ii) BoF and SVM method.

1) Classification using Contour Comparisons

In this method, the contour properties such as solidity and area of the contour are extracted and compared with already assumed values to determine whether the vehicle is LTV, HTV or MTV. The contours are basically all connected pixels in an image. Once the background subtraction is performed and foreground objects are detected; then contours of these foreground objects are detected.

We have used OpenCV's find Contours () method to detect the contours of the foreground object. The method gives a list of all the contours in the image. It is necessary to choose the biggest contour in the image. Hence, a minimum limit on the width and height of the contours is defined so that bigger contours can be chosen. Once these contours have been selected, a number of properties of these contours are extracted which can later be used to classify the vehicles. Some of these properties are area, solidity, aspect ratio etc. Particularly, areas of contours are focused which are compared to the assumed values of the vehicles.

First of all, a bounding box is drawn over the contour and its centroid are found which when intersects with the imaginary line, the vehicles are detected and classification algorithm is triggered to classify the vehicles.

The following values are assumed:

If the area is between 500 and 8000, the vehicle is classified as LTV.

Classification of MTV and HTV is two-step process:

1. If area is between 500 and 125000 the vehicle could be either MTV or HTV depending upon the height.
2. If height of the vehicle is about its width (width-30 to width+30) it is classified as MTV else it is classified as HTV.

2) BoF and SVM method

This is another method of classification which uses BoF along with SVM algorithm to classify the vehicles. The BoF is based on the bag of words algorithm where each word corresponds to each feature of the image saved in the bag. The SVM classifier is created with four classes which are LTV, MTV, HTV and Others. The LTV class contains the features of the small vehicles, MTV contains the medium vehicles and HTV contains the big vehicles, whilst others contain the vehicles that don't fall into any class such as bikes etc. The images of all vehicles are placed in the particular folders as defined above and the SVM classifier is trained over these folders. Once the classifier is trained, it gives a file called BoF.pkl which is used in the code to test the classifier. The system captures the SIFT features of the foreground objects and compares them with the already trained SVM and classifies the vehicles into defined categories.

V. SIMULATION RESULTS

The proposed system is implemented on python, using the OpenCV bindings. The traffic camera footages from variety of sources have been used and have found the implementation to be effective. Basically, it is divided into

two parts: detection and classification. Five sample videos of different traffic scenes are used for comparison of the system results with true values. Fig. 5 presents the comparison of total vehicles counted by both CC and BoF and SVM methods with the true values. It is observable that CC method is showing smallest discrepancy with counts of ground truth related to video 2 and counts related to video 1 are showing highest discrepancy. BoF and SVM is depicting smallest variation with ground truth count related to video 3 and showing highest variation with counts related to video 1. Fig. 6 to Fig. 10 compares the simulation results for the classification. The proposed system classifies the detected vehicles into three categories i.e. LTV, MTV and HTV using two different methods. First one deploys the contour area and assumed area of vehicles for the comparison and second method deploys the BoF and SVM for classification. From fig. 6 to fig. 10, it is evident that in all the videos; CC method classifies LTV, MTV and HTV very well as compared to BoF and SVM method except for video 3. The values of CC method are closer to ground truth values than BoF and SVM method. The results concerning the classification errors of the both methods which are employed in the experiments for different video sequences of road traffic are shown in Table I. From the table it is apparent that classification errors of CC method are lower as compared to BoF and SVM method in all videos.

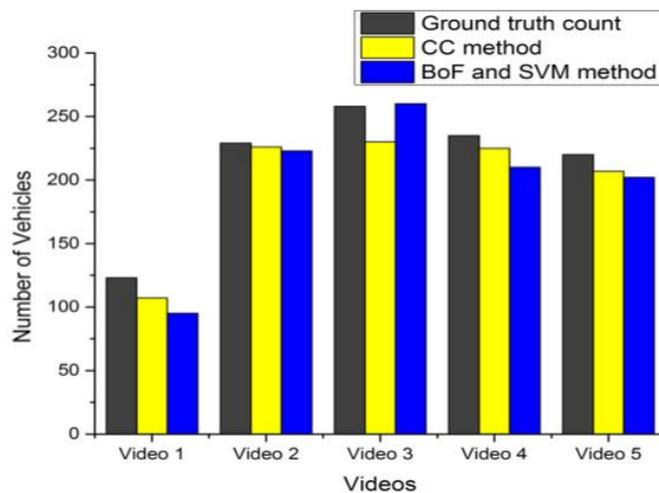


Fig.5. Comparison of original count and system count values of five videos.

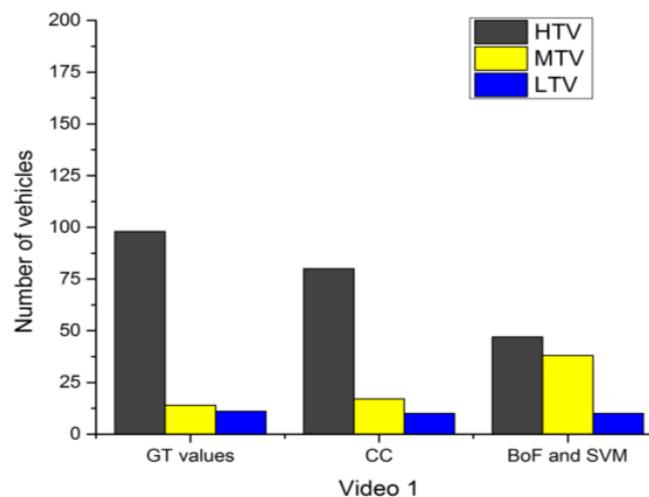


Fig.6. Vehicle classification comparisons between original count, CC, SVM and BoF methods in video

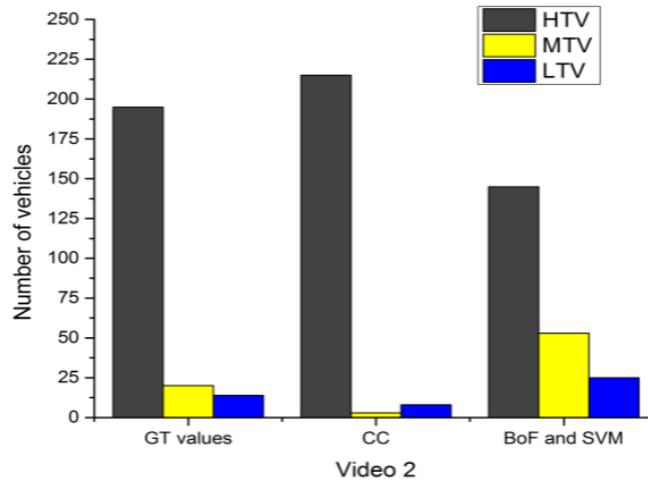


Fig.7. Vehicle classification comparisons between original count, CC, SVM and BoF methods in video.

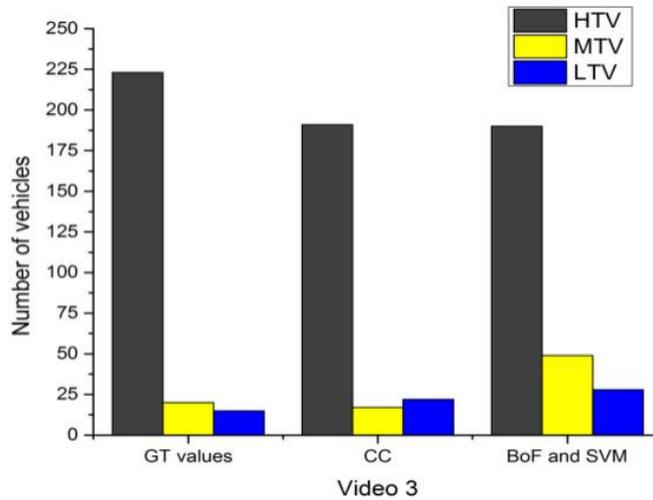


Fig.8. Vehicle classification comparisons between original count, CC, SVM and BoF methods in video.

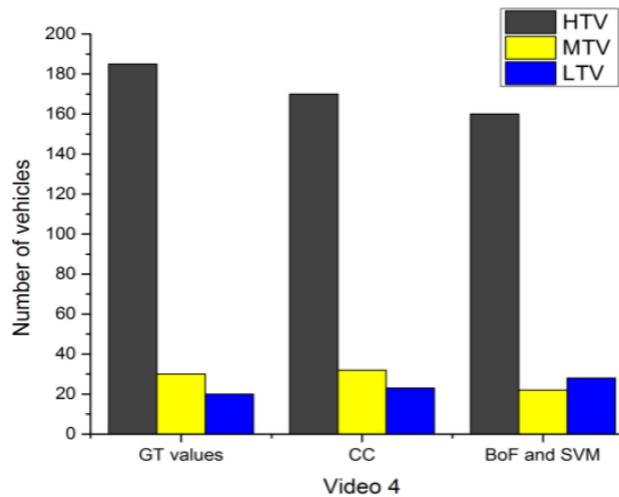


Fig.9. Vehicle classification comparisons between original count, CC, SVM and BoF methods in video.

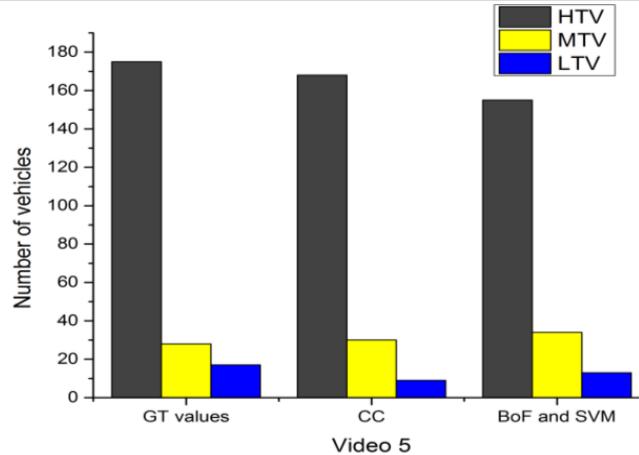


Fig.10. Vehicle classification comparisons between original count, CC, SVM and BoF methods in video.

Table 1. Vehicle classification errors of CC, SVM and BoF methods

	Ground Truth Count	Classification Error % (CC)	Classification Error % (BoF and SVM)
Video 1	123	13	22.76
Video 2	229	1.3	2.6
Video 3	258	10.85	0.7
Video 4	235	4.25	10.6
Video 5	220	5.9	8.18

VI. CONCLUSION

The proposed solution is implemented on python, using the OpenCV bindings. The traffic camera footages from variety of sources are in implementation. A simple interface is developed for the user to select the region of interest to be analyzed and then image processing techniques are applied to calculate vehicle count and classified the vehicles using machine learning algorithms. From experiments it is apparent that CC method outperforms than BoF and SVM method in all results and gives more close classification results to the ground truth values. Currently proposed system works with already captured videos but it can be modified to be used for processing live video streams [4] by adding microcontrollers. One of the limitations of the system is that it is not efficient at detection of occlusion of the vehicles which affects the accuracy of the counting as well as classification. This problem could be solved by introducing the second level feature classification such as the classification on the bases of color. Another limitation of the current system is that it needs human supervision for defining the region of interest. The user has to define an imaginary line where centroid of the contours intersects for the counting of vehicles hence the accuracy is dependent on the judgment of the human supervisor. Furthermore, the camera angle also affects the system hence camera calibration techniques could be used for the detection of the lane for the better view of the road and increasing the efficiency. The system is not capable of detection of vehicles in the night as it needs the foreground objects to be visible for extraction of contour properties as well as features for the classification using SIFT features [31]. The system could also be improved for better accuracy using the more sophisticated image segmentation and artificial intelligence operations.

VII. REFERENCES

- [1] S.-Y. Cheung, and P.P. Varaiya, "Traffic surveillance by wireless sensor networks: Final report", PhD diss., University of California at Berkeley, 2006.
- [2] S. Oh, S. Ritchie, and C. Oh, "Real-time traffic measurement from single loop inductive signatures", Transportation Research Record: Journal of the Transportation Research Board, (1804), pp. 98-106, 2002.

- [3] B. Coifman, "Vehicle level evaluation of loop detectors and the remote traffic microwave sensor", *Journal of transportation engineering*, vol. 132, no.3, pp. 213-226, 2006.
- [4] M. Tursun, and G. Amrulla, "A video based real-time vehicle counting system using optimized virtual loop method", *IEEE 8th International workshop on Systems Signal Processing and their Applications (WoSSPA)*, 2013.
- [5] M. Lei, D. Lefloch, P. Gouton, K. Madani, "A videobased real-time vehicle counting system using adaptive background method", *IEEE International conference on Signal Image Technology and Internet Based Systems (SITIS'08)*, pp. 523-528, 2008.
- [6] N.C. Mithun, N.U. Rashid, and S.M. Rahman, "Detection and classification of vehicles from video using multiple time-spatial images", *IEEE Transactions on Intelligent Transportation Systems*, vol 13, no 3, p. 1215-1225, 2012.
- [7] R.T. Collins, et al., "A system for video surveillance and monitoring", *VASM final Report*, Robotics Institute, Carnegie Mellon University, 2000, pp.1-68.
- [8] G. Yang, "Video Vehicle Detection Based on SelfAdaptive Background Update", *Journal of Nanjing Institute of Technology (Natural Science Edition)*, vol 2, p. 13, 2012.
- [9] F. Liu, and H. Koenig, "A survey of video encryption algorithms", *computers & security*, vol. 29, no 1, pp. 3-15, 2010.
- [10] E. Bas, A.M. Tekalp, and F.S. Salman, "Automatic vehicle counting from video for traffic flow analysis", *IEEE Intelligent Vehicles Symposium*, 2007.
- [11] H. Rabiou, "Vehicle detection and classification for cluttered urban intersection", *International Journal of Computer Science, Engineering and Applications*, vol 3, no 1, p. 37, 2013.
- [12] M. Seki, H. Fujiwara, and K. Sumi, "A robust background subtraction method for changing background", *Fifth IEEE Workshop on Applications of Computer Vision*, 2000.
- [13] J.J. Gibson, "The perception of the visual world", 1950.
- [14] K. Wu, et al., "Overview of video-based vehicle detection technologies", 2011.
- [15] N. Friedman, and S. Russell, "Image segmentation in video sequences: A probabilistic approach", *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*, 1997, Morgan Kaufmann Publishers Inc.
- [16] C. Stauffer, and W.E.L. Grimson, "Learning patterns of activity using real-time tracking", *IEEE Transactions on pattern analysis and machine intelligence*, 2000. Vol 22, no 8, pp. 747-757, 2000.
- [17] C. Stauffer, and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking", *IEEE Computer Society Conference Computer Vision and Pattern Recognition*, 1999.
- [18] A. Elgammal, D. Harwood, and L. Davis, "Nonparametric model for background subtraction", *European conference on computer vision*, Springer, 2000.
- [19] S. O'Hara, and B.A. Draper, "Introduction to the bag of features paradigm for image classification and retrieval", *arXiv preprint arXiv:1101.3354*, 2011.
- [20] D.G. Lowe, "Object recognition from local scale-invariant features", *Computer vision, Seventh IEEE international conference*, 1999.
- [21] D.G. Lowe, "Distinctive image features from scaleinvariant keypoints", *International journal of computer vision*, vol 60, no 2, pp. 91-110, 2004.
- [22] V. Ramakrishnan, A. K. Prabhavathy, and J. Devishree, "A survey on vehicle detection techniques in aerial surveillance", *International Journal of Computer Applications*, vol 55, no 18, 2012.
- [23] B. Pawar, V.T. Humbe, L. Kundani, "Morphology Based Moving Vehicle Detection". *International Conference on Big Data Analytics and computational Intelligence (ICBDACI)*, pp. 217-223, 2017.
- [24] R.H. Pena-Gonzalez, M.A Nuno-Magada, "Computer vision based real-time vehicle tracking and classification system", *IEEE 57th International Midwest Symposium on Circuits and systems (MWSCAS)*, pp. 679-682, 2014.
- [25] A. Suryatali, V.B. Dharmadhikari, "Computer Vision Based Vehicle Detection for Toll Collection System

- Using Embedded Linux”, International Conference on Circuit, Power and Computing Technologies (ICCPCT), pp. 1-7, 2015.
- [26] N. Seenouvang, U. Watchareeruetai, C. Nuthong, K. Khongsomboon, “A Computer Vision Based Vehicle Detection and Counting System”, IEEE 8th International conference on Knowledge and Smart Technology (KST), pp.224-227, 2016.
- [27] A.B. Godbehere, A. Matsukawa, and K. Goldberg, “Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation”, IEEE, American Control Conference (ACC), pp. 4305-4312, 2012.
- [28] P. KaewTraKulPong, and R. Bowden, “An improved adaptive background mixture model for real-time tracking with shadow detection”, Video-based surveillance systems, Springer. pp. 135-144, 2002.
- [29] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction”, 17th International Conference on Pattern Recognition (ICPR), 2004.
- [30] Z. Zivkovic, and F. van der Heijden, “Efficient adaptive density estimation per image pixel for the task of background subtraction”, Pattern recognition letters, vol 27, no 7, pp. 773-780, 2006.
- [31] K. Robert, “Night-time traffic surveillance: A robust framework for multi-vehicle detection, classification and tracking”, Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS'09), pp. 1-6, 2009.