Video Monitoring System at Fuel Stations

Gauri Sabnis
Dept of Computer Engineering
Vishwakarma Institute of Technology
Pune, India

S. R Shinde
Dept of Computer Engineering
Vishwakarma Institute of Technology
Pune, India

Abstract—Computer vision is a scientific and digital field which deals with how computers are made to gain high-level understanding from digital images or videos. It is used to automate tasks involving digital images and videos. Computer vision mainly includes acquiring, processing, analyzing and understanding digital images and extraction of high-level data from the real world videos and images in order to make in the forms of decisions. The objective is to use computer vision techniques to monitor a public space i.e. fuel stations in terms of surveillance as well as safety and security. In this proposed system a live stream feed from a camera placed at a fuel station is taken and perform video processing and computer vision techniques to give real time results in the form of quantitative information about the vehicles, people present and general monitoring of the fuel station. The data collected overtime through the system can be analyzed to implement further improvements for betterment of the business.

Keywords— Object recognition, image processing, object detection, neural networks

I. INTRODUCTION

Computer Vision can be used to automate tasks involving digital images and videos. Computer vision mainly includes acquiring, processing, analyzing and understanding digital images and extraction of high-level data from the real world videos and images in order to make in the forms of decisions. Video Analytics is a part of computer vision wherein videos are acquired and analyzed using various techniques to get the desired output. The goal we are trying to accomplish is taking live stream feed from a camera placed at a fuel station and perform video processing techniques to give results in the form of quantitative information about the vehicles and general monitoring of the fuel station. This data collected overtime can be analyzed to implement improvements for betterment of the business. Analysis of the video footage or live stream content produced at the fuel stations enables to organize, share and analyze any insights gained from the data to make smarter and better business decisions for the fuel station business. The system focuses on usage of existing infrastructure i.e cameras already installed at the fuel stations to gain insights into the business using different use cases. This being a public place there exist problems including safety security surveillance at the fuel stations so the system will be able to monitor the area. The system will also help in the need for an intelligent video system for profile creation of the visiting customers at the station. Data extracted from the analysis can be stored in a suitable database and applied data mining to gain insights and better the existing operations. Real time processing of the video content is a constraint. So, time constraint becomes of vital importance here. A time lag to a certain degree can be acceptable. Fast processing of the live feed/video is required. Hence, requires graphical processing units available. Many occlusions occurring especially in heavy traffic reduce the accuracy of object detection greatly. The paper focuses on development of a system useful for general monitoring and surveillance at a fuel station. The organization of the paper is as; some of the techniques that are used for multiple use cases in recent years are introduced and analyzed in the literature survey section II. The proposed methods discussed in section III followed by the results and conclusions in section IV.

II. LITERATURE SURVEY

Many works and methods related to the various use case already exist, which have numerous traditional methods based on image processing techniques. Literature survey is done based on each individual use case.

Object Detection systems: Automatic detection of objects in the image/video can be a challenging problem. There are various techniques for multiple object detections like humans and vehicle detection, ranging from generic algorithms to specific high level architectures especially developed to deal with human body detection. The most common challenges for detection as well as tracking are collision of objects on image and occlusion of objects. So, it is important that the algorithm re-identifies target in other camera angles, after or the time it is temporarily totally or partially occluded by any other thing in the scene. The task becomes more difficult when two or more people (tracked) occlude each other in front of the camera and it causes identification problem as is recognized in [1]. The main notion of multi-object tracking system is to include a priori information about the object of interest like chair, bike, windows etc. In case of dealing with human object, information about human shape is used. Similar process is used with the detection of vehicles. In cases of surveillance number of cameras can be present at the site thus multi camera solution is also discussed in [2] and for a site with a dense camera network useful for surveillance is discussed in [8]. Though, in this system the method used for object detection is Haar Cascade classifier based on the paper written and proposed by P. Viola and M. Jones [18]. To sum up the method in simple terms, it is a machine learning algorithm where a cascade function is being trained on a lot of positive and negative images. Positive images are defined as images having the object to be recognized and negative images are the ones not containing the object of recognition. The cascade function will be used to detect objects. Haar features are single values obtained by subtracting the sum of pixel intensities under the black rectangles. The convolution

here dictates how the sum of pixels in the white and black part interact mostly they are differentiated to produce a single value. The algorithm for calculating this is called integral functions. Next we have feature selection in which top features that are important are chosen. Using these features different stages of these features are created called cascade features used to identify and recognize the object. That is why the method is called Haar Cascade classifier. Method for human tracking under unpredictable trajectories is presented in [10]. The technique uses an omega-shaped descriptor. The tracking uses particle filter with linear filter to determine the next position of a tracked person. The error produced by the particle filter in next coming frames is handled by the improved Viola-Jones and HOG feature based SVM detector. When either a collision or occlusion event is detected, the particle filter (used to track every person in the scene) is disabled. The regions of scene in which the lost target could present are defined by elliptic blobs. Newly appeared target is compared against the lost target using color-histogram representation. The method proposed in [11] deals with detection and tracking. The detection includes face and eye detection.

Systems based on vehicle detection and traffic analysis with vehicle registration number plate recognition: In the traffic surveillance, the division of road information brings advantages, e.g. it enables to automatically crop regions of traffic analysis. So, it speeds up flow in the videos and helps with the detection of driving or traffic monitoring and violations. Road segmentation improves information in the videos of traffic. In urban scenes the following challenges are faced: the urban scenes exhibit more cluttered scenarios; these consist of many vehicles stopped on the roads; and are not in parallel lines, they also intersect under various angles; and finally, pedestrians can complicate the challenge further with the road boundaries since they can cause errors [6]. The authors of [16] talk about several important challenges of automatic vehicle recognition in video systems that have multiple camera settings and propose new technologies for such systems. The focus is on two main issues for vehicle recognition: automatic vehicle recognition and vehicle registration number plate recognition. The automatic vehicle recognition deals with such challenges as multiple-feature fusion, different camera properties for region extraction and on profit object detection/recognition. An idea for Automatic Vehicle Recognition, multiple cameras for the video application presented in this paper is on the proposed adaptive vehicle registration number plate detection technique, connected component analysis and level-based region comparison algorithm. Also, for improving LPR accuracy, the paper presents LPR method on neural network to identify alphanumeric characters of vehicle registration number plates. Experimental results indicate simplicity and effectiveness of the proposed algorithms.

Object counting/tracking systems: Counting the number of objects in a video is a challenging task in computer vision. From the video perspective, the most issues are related with people and vehicle counting in the sense of traffic is dense. Counting by detection and tracking is common [4]. But, it requires explicit object detection and modeling, and occlusion handling. The authors of [17] propose the counting method for arbitrary objects in the scene. The method automatically counts the area of interest by the motion flow of objects in the area. The other paper [18] presents an efficient self-learning people counting system that is innovative. It can accurately count the number of people in a region of interest. To detect the pedestrians bag-of-features model is used. It provides a good distinction between standing or slowly moving pedestrians, and the background. The proposed system had an ability of automatic pedestrian and non-pedestrian samples selection in order to update the classifier. It has ability to a real-time adaptation to the specific scene. Experimental results proved the robustness and high accuracy of the system.

III.METHODOLOGY

The different use cases to be considered in the system are as follows

1. Object Counting: Object counting implies number of vehicles present at the fuel stations in different regions. To estimate crowd density at fuel stations individual stalls gives an estimate of peak sales times at the fuel station.

2. Object Classification: Object classification implies classification of types of vehicles into various categories according to their fuel consumption. Smaller consumption vehicles such as two wheelers or four wheelers like cars and larger consumers such as lorries or trucks can be differentiated for the betterment of the business. The consumers can be identified and analysis can be performed.

3. Detection of stalls available or unavailable: Identification of fuel station stall area and detect if vehicle is present or not. Count the number of vehicles if present simultaneously taking services

4. Detection of Humans: Detection of attendants present for each customer for good customer satisfaction.

5. Recognition of vehicle registration number plates: Detection and recognition of vehicle registration number plates for each vehicle for security purposes.

6. Dashboard Creation: Creation of dashboard to display real time parameter values mentioned above for better understanding and management.

Algorithm:

Input: Live streaming video

1. Initialization
2. Define a function that takes the hardcoded coordinates as arguments and marks the stalls on live video and returns stall _boundary coordinates
3. Define a function to detect each class of vehicle using haar cascade classifier and returns vehicle type as well as system in and out time of vehicle when enters stall
4. Define function to Detect dispenser unit manager. Detection of human every 1 minute. Return system time of detection

5. Define function to detect if vehicle in and out time is less than 30 seconds. If yes then business loss else business profit
6. Define function to check availability or unavailability of stall by checking if the vehicle lies in stall_boundary

Output: Live dashboard and excel files of per minute data

Access to live streaming is provided as input .The stalls on the fuel station are marked on the video using either constant coordinates or drawn over on the video. The incoming vehicles are first classified into vehicle categories like cars, bikes, truck, lorries etc total 8 categories. For certain vehicle categories like auto rickshaws that are pertaining to India, separate training for identification of such vehicles is required. After detection of incoming vehicle, as it enters the boundary of a stall, the stall is shown busy. If a stall has no vehicles currently in it, it is shown as available and if it has one or more vehicles present in it, it is shown as unavailable. After a vehicle comes to halt in a stall, the availability of an attendant is detected. The time of the arrival of the vehicle and the duration of it until it leaves is noted. The salesman is detected in the position every 1 min if he is present or not. For all the dispensing unit stall areas, the vehicle detected is counted according to each type. The in and out time of the vehicle is noted and if a vehicle is at the dispensing unit for more than 30 seconds then it's classified as business profit otherwise as business loss. A screenshot of the vehicle registration number plate of the vehicle is taken. The vehicle registration number plate numbers are recognized and displayed. A count of vehicles in each stall per day is maintained. All files containing data of the following for one day is saved: Vehicle type detected, vehicle in and out time, business profit/loss, and vehicle count and dispenser salesman present on each 1 min interval at Dispenser position. Dashboard is created which displays the live stream for CCTV cameras as well as real time analysis of the video content is displayed. So features like stall availability, count of the number of vehicles per stall, Attendant present per stall, Vehicle registration number plate number of the currently present vehicle at each stall. The last step is storing all of the features in a suitable database

Fig. 1. Activity diagram for the system

The main components are:

Vision for object detection and tracking/counting: vehicle detection technique used for Detection and classification was Haar cascade classifier. Haar Cascade Classifier is a method for detecting the object, also referred to as Viola Jones method due to its authors Paul Viola and Michael Jones first for face detection. This method has 4 sub methods for detecting any object, such as Haar-like feature, integral image, AdaBoost learning and Cascade Classifier. Haar-like feature which is the difference of pixel intensities is high speed computation that depends on the number of pixels in the rectangle feature selected and not depends on each pixel value in the image. For object detection value calculation, Haar like feature value is first calculated using integral image. The integral image value is obtained by summing up value of previous indices, started by left top until right bottom. Strong classifier made by AdaBoost selecting strong features can detect object in stages on a cascade. All sub-windows are scanned for this criteria on each step, After that, a sub-window containing the desired object is used as an input for the next stage of filtering with more specific criteria until a sub-window which is predicted as a car is obtained. The sub-windows not containing the object is considered as background and separated. Object Tracking Object tracking is used to obtain the specific position (x,y) coordinates of object in the frame to be compared with the previous positions of tracked objects, but, newly discovered positions or positions already not present on the list of tracked object positions is added as a position (x,y) of a newly detected object. If the new position is present in the list of coordinates of previous tracked objects, it is declared as a new position of a recognized object. Object counting every passing vehicle object inside ROI (Region of Interest) was tracked based on its position and would be compared with the list of tracked object positions. For a new position or position not including in the list of tracked objects, it is checked to be at a Euclidian distance 50 or less. If the new position was included in the list of positions of previous tracked objects, it means the position had already been counted as a recognized vehicle.

Vehicle registration number Plate Recognition: Automatic number-plate recognition (ANPR) is a technology that includes and combines optical character recognition on images to recognize vehicle registration plates to create vehicle data. It can use existing closed-circuit television footage, road rule enforcement cameras installed by government, or cameras specifically designed for ALPR (Automatic License Plate Recognition). Automatic vehicle registration number plate recognition has two most important technological issues that is the quality of the vehicle registration number plate recognition software given its applied recognition algorithms should be strong, and the quality of the image acquisition technology, the camera and the illumination of the scene also the camera angle (which should be between 0 to 30 degrees), and the distance from the vehicle etc. The important factor is the vehicle registration number plate recognition software part. The complexity of the recognition software part, the quality of the vehicle registration number plate recognition algorithms, 20 the mathematical application and the years of experience behind
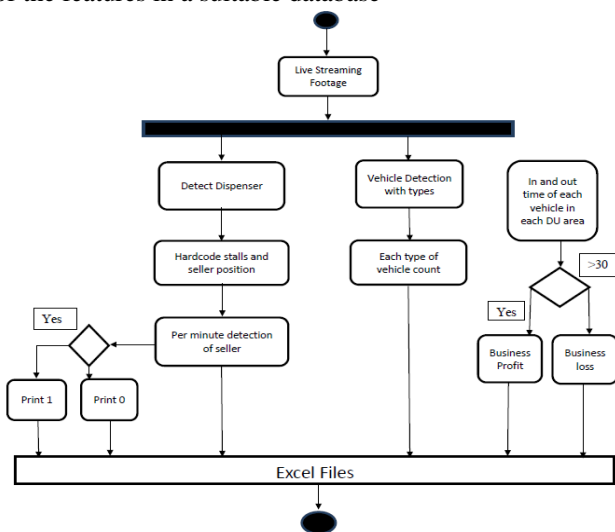
it determines the capacities of the recognition software. The better the algorithms performing are, the greater the quality of the recognition software is: the highest recognition accuracy it has, the fastest processing speed it has, the most type of plates it can handle, the widest range of picture quality it can handle, the most tolerant against distortions of input data it is. In 2017, OpenALPR reported accuracy rates for their commercial software in the range of 95-98% on a public image benchmark.

The other method tested for vehicle registration number plate recognition is Pytesseract OCR (Optical Character Recognition). Pytesseract reads the text present in any document like jpg, png, tiff, bmp etc. Pytesseract is a python tool for OCR and can recognize in over 100 languages. Vehicle registration number plate recognition using this method needs steps such as :

1. Image Preprocessing
2. Plate Localization and Extraction
3. Character Segmentation
4. Character Recognition

The above method is based on the works of Andrew S. Agbemenu, Jepthah Yankey and Ernest O. Addo [19]

IV. RESULTS AND CONCLUSIONS

The detection and tracking of the 8 categories of vehicles was done by using Haar cascades. The 8 categories namely cars, trucks, tractors, tankers, buses, two wheelers, tempo, and auto rickshaw each category classifier was built separately by using 200 positive and 200 negative images of each category. Apart from the accuracy obtained after training the model with positive and negative images some test images have been used.

Test Cases:

The test cases results for each of the algorithms used are as follows:

Vehicle and Human Detection and tracking: Detection using and different algorithms were tried and the results shown are:

Table.1. Comparison of different techniques of object detection

Techniques used	Maximum object detection confidence achieved	Occlusion support	Relative detection of objects
YOLO (You Only Look Once)	0.97	less	Problem while detecting moving objects. Object tracking failure
R CNN	0.83	more	Slow detection of objects
Haar cascades classifier	0.95	more	Highest number of accurate detections

The above table shows that even though YOLO (You Only Look Once) is generally used for real time object detection, in this case tested on a number of test cases it can be seen haar classifier performing more accurately. Hence for

object detection and tracking Haar cascade classifiers are used and trained.

The Haar cascade model trained was on cascade taking 200 positive and 200 negative images of all categories. The positive images are objects on roads. Traincascade was used for training on opencv 3.4. The training is done in 10 stages till best features are selected. For car xml training 7 best features were selected.
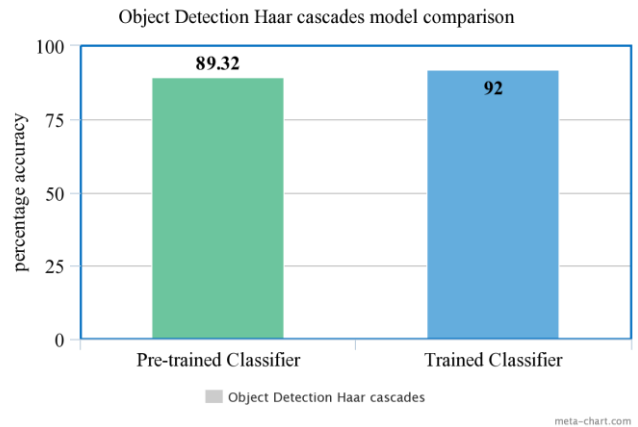

Fig.2. Training of the xml classifier for cars model accuracy comparison

Vehicle registration number Plate recognition: For test cases 7 images having vehicle registration number plates from various angles. The two approaches used are Open ALPR and pytesseract for OCR.

Table.2. Test cases for OpenALPR vs. Pytesseract OCR

Images for testing	OpenALPR	Pytesseract OCR
HR26DK8337	Partial characters recognized	Full characters recognized
MH 20 EJ 0364	Partial characters recognized	Partial characters recognized
TN 38 BR 3036	Partial characters recognized	Full characters recognized
TN 99 F 2378	Partial characters recognized	Full characters recognized
MH-12-TC-605-K	Partial characters recognized	Full characters recognized
MH12BG7237	Partial characters recognized	Partial characters recognized
KL 65 K 7111	Partial characters recognized	Full characters recognized

As shown in the test cases we can see that for object detection and human detection the approach used Haar Cascade will give us more accurate results in live feed situation. Haar cascade method still fails considerably when

collision or occlusion appears. Compared to other deep learning approaches the Haar cascade method is useful for quick recognition but it takes a lot of time for training.

The method used for vehicle registration number plate recognition were two methods tried firstly using Open ALPR is an open source automatic vehicle registration number plate recognition library that is commercially distributed also and written in C++ also has bindings for python. The second approach is finding contours in images finding the number plate and using Pytesseract library to recognize characters using Optical recognition.
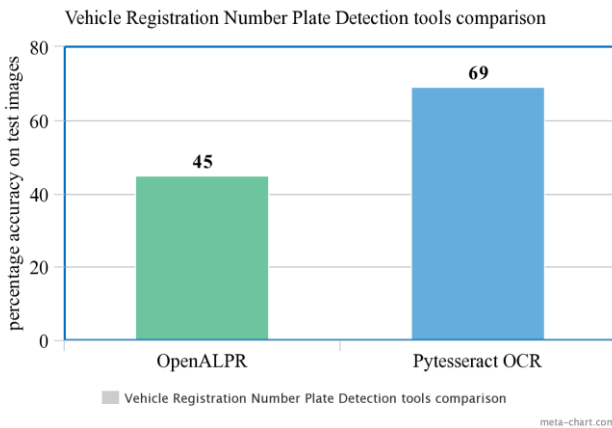

Fig.3. OpenALPR and Pytesseract OCR comparison

The steps included in the recognition of vehicle registration number plate are:
1. Grayscale conversion
2. Canny edge detection
3. Contour detection
4. Vehicle registration number plate detection

Grayscale coverts RGB image into grayscale for better operation on the image. The canny edge detection detects the edges in the images. With the contour detection the number plate is detected [20] with respect to given number plate dimensions and colour contrast. Finally when the number plate is detected it is then fed to the Pytesseract OCR module for character segmentation.

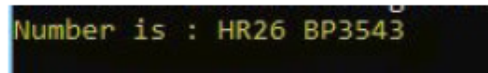Fig.4. Vehicle registration number plate detection

Fig.5. OCR output

The system on Video content analysis of video footage is an intelligent system that can be used for multiple use cases and not just the ones mentioned. The use cases can provide with many opportunities to improve upon strategies using the long term data from this proposed system. As mentioned earlier this can be significant to a lot of domains including security, safety and precautions, and customer satisfaction. Thus there is further scope of improvement using other newly developed technologies as well.

The future scope of the system includes implementation of other use cases that can enhance the system useful for safety and security of the customers as well as the operations of the station like detection of anomalous behaviors, technical difficulties, and implementation of rules pertaining to safety, detection of drivers and recognition of drivers for more safety against criminal activities.

REFERENCES

[1] Frejlichowski, D., Forczmański, P., Nowosielski, A., Gościewska, K., Hofman, R.: SmartMonitor: An Approach to Simple, Intelligent and Affordable Visual Surveillance System. In: Bolc, L. et al. (eds.) ICCVG 2012. LNCS, vol. 7594, pp. 726–734. Springer, Heidelberg, 2012.

[2] Yunbo Rao, Automatic vehicle recognition in multiple cameras for video surveillance, Springer-Verlag Berlin Heidelberg 2014, Vis Comput DOI 10.1007/s00371-013-0917-y

[3] MichałZabłocki, KatarzynaGościewska,DariuszFrejlichowski, RadosławHofman, Intelligent video surveillance systems for public spaces –a survey, Journal of Theoretical and Applied Computer Science Vol. 8, No. 4, 2014,

[4] Cancela, B., Ortega, M., Penedo, M.: Multiple human tracking system for unpredictable trajectories. Machine Vision and Applications, 25(2), 511-527, 2014.

[5] Tathe, S., Narote, S.: Real-time human detection and tracking. 2013 Annual IEEE India Conference (INDICON), pp. 1-5, 2013.

[6] Kushwaha, A., Sharma, C., Khare, M., Srivastava, R., Khare, A.: Automatic multiple human detection and tracking for visual surveillance system. 2012 International Conference on Informatics, Electronics Vision (ICIEV), pp. 326-331, 2012

[7] Andersson, M., Gustafsson, F., St-Laurent, L., Prevost, D.: Recognition of Anomalous Motion Patterns in Urban Surveillance. IEEE Journal of Selected Topics in Signal Processing,. 7(1), 102-110, 2013.

[8] Calavia, L., Baladrón, C., Aguiar, J. M., Carro, B., Esguevillas, A. S.: A Semantic Autonomous Video Surveillance System for Dense Camera Networks in Smart Cities. Sensors, 10407- 10429, 2012.

[9] Lim, M. K., Tang, S., Chan, C. S. iSurveillance: Intelligent framework for multiple events detection in surveillance videos. Expert Systems with Applications, 41(10), 4704-4715, 2014.

[10] Lee, S., Nevatia, R.: Hierarchical abnormal event detection by real time and semi-real time multi-tasking video surveillance system. Machine Vision and Applications, 25(1), 133-143,2014. *Intelligent video surveillance systems for public spaces – a survey 27*

[11] Su, H., Yang, H., Zheng, S., Fan, Y., Wei, S.: The Large-Scale Crowd Behavior Perception Based on Spatio-Temporal Viscous Fluid Field. IEEE Transactions on Information Forensics and Security, 8(10), pp. 1575-1589, 2013.

[12] Onal, I., Kardas, K., Rezaeitabar, Y., Bayram, U., Bal, M., Ulusoy, I., Cicekli, N.: A framework for detecting complex events in surveillance videos. 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 1-6, 2013.

[13] Santos, M., Linder, M., Schnitman, L., Nunes, U., Oliveira, L.: Learning to segment roads for traffic analysis in urban images. 2013 IEEE Intelligent Vehicles Symposium (IV), pp. 527-532, 2013.

[14] Rao, Y.: Automatic vehicle recognition in multiple cameras for video surveillance. The Visual Computer, 1-10, 2014.

[15] Zhou, Y., Luo, J.: A practical method for counting arbitrary target objects in arbitrary scenes. 2013 IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6, 2013.

[16] Li, J., Huang, L., Liu, C.: An efficient self-learning people counting system. 2011 First Asian Conference on Pattern Recognition (ACPR), pp. 125-129, 2011.

[17] Paul Viola, Michael Jones: Rapid Object Detection using a Boosted cascade of Simple Features. Accepted conference on computer vision and pattern recognition, 2001

[18] Andrew S. Agbemenu, Jepthah Yankey, Ernest O. Addo: An Automatic Number Plate Recognition System using OpenCV and Tesseract OCR Engine. International Journal of Computer Applications (0975 - 8887) Volume 180 - No.43, May 2018