AUTOMATED TRAFFIC MONITORING FOR COMPLEX ROAD CONDITIONS

Opatha, R. K;Peiris, Akila;Gamini, D. D. A.;Edirisuriya, Ananda ;Athuraliya, C. D. ;Jayasooriya, Isuru;

;

IDRC Grant/ Subvention du CRDI: 108008-001-Leveraging Mobile Network Big Data for Developmental Policy

Annex 16: Automated Traffic Monitoring for Complex Road Conditions

R.K. Opatha, Akila Peiris, D.D.A. Gamini, Ananda Edirisuriya, C.D. Athuraliya, Isuru Jayasooriya

INTRODUCTION

Traffic management is one of the major pressing issues in any urban setting. Situation is severe especially in developing countries due to the lack of infrastructure and up to date regulations. Authorities report that 200,000 vehicles come to Colombo, Sri Lanka daily. This traffic is monitored by 105 Closed Circuit Television (CCTV) cameras installed at main intersections in the Colombo City and footage transmitted to a central control room with 28 screens [1]. But the actual traffic monitoring is mostly done manually by human experts. Even though this can be very accurate it is not scalable. Continuously monitoring multiple screens 24x7 can be a very tedious task for even a team of human experts. And as we increase the number of CCTV installations to other areas the monitoring team will need expansion proportionally.

As a solution to aforementioned problem many developed countries have automated the traffic monitoring process up to possible extents with the help of recent computer vision and machine learning techniques. CCTV monitoring systems have been developed and installed with the infrastructure to automatically detect different interested entities such as vehicles, pedestrians and traffic violations from CCTV footage. With the recent advancements in computer vision and machine learning techniques, these systems have become highly effective in certain cases where you have well structured vehicle traffic such as highways. But these systems still struggle in handling complex and irregular traffic conditions.

In this work we seek to develop a solution that works in complex conditions by applying recent computer vision and machine learning as well as deep learning techniques. We break down the problem into different sub-tasks such as vehicle detection, vehicle tracking, vehicle recognition and combine each process into a one pipeline that can be used for traffic monitoring. We applied multiple existing techniques for each aforementioned sub-task and evaluated them to select best performing technique for each stage of the pipeline. Implementing the final pipeline involved improving and aggregating existing techniques. The results demonstrate the potential of these techniques for automated traffic monitoring.

Extending above work we also explored the possibility of detecting traffic violations from CCTV footage. This task is inherently more challenging than the previous tasks and there are no widely accepted frameworks or evaluation metrics for traffic violation detection. Nevertheless our early work for few violation types yielded modest results and exhibit the possibility of using CCTV footage for traffic violation detection.

As it was mentioned earlier, an automated traffic monitoring system involves multiple stages such as foreground estimation, object detection, object tracking and object recognition. Most of the existing systems do not include all these functionalities within a one system. And also most of these systems are commercial and proprietary. We can see multiple efforts in relevant literature for different stages of this pipeline.

Contour based classification is a method of encountering the edge of the silhouette into an account. This mechanism is used to deal with the problem of occlusion between vehicles. The idea proposed by [2] had proposed side view contour based classification process to resolution of occlusions between two different vehicles to separate those two vehicles if those occlusions are not critical. The system proposed by [3], had offered techniques using neural network classification is being handled in two steps using two different neural networks considering the geometrical parameters resulting an accuracy of 69% using 100 vehicles. The system proposed by [4] also performed a similar work using machine learning concepts by instead tracking regions and using the fact that all motion occurs in the ground plane to detect, track and classify vehicles. Unfortunately, this mechanism did not address the vehicle shadows represent in the videos. The system proposed by [5] had a vehicle classification on Indian roads using procedures like SIFT descriptors and SVM. They had used different kernel types for the experimental purposes. Linear, Quadratic and Radial Based Functions are being used as the kernel types with the support vector classifier resulting 78.54%, 81% and 90% of accuracy based on the combination.

The system proposed by [6] had an on-road vehicle detection procedure using rear views of the vehicles. They have used feature extraction and classification to detect the vehicles. They have specifically proposed the Gabor filter for feature extraction and Support Vector Machines for classification and vehicle detection. The system proposed by [7] had offered the concept of Bayesian network for vehicle classification approach. Their experiments resulted an accuracy level of 95.7% using 177 vehicles in four distinct categories.

Overall, with the technological evolution object detection, object tracking and object classification has become the most researched area in Computer Science Systems. This provides the capability of actively involved in the process, provided with the personalized, low cost features to improve the overall experience and most importantly the benefits in the traffic management domain in Sri Lanka.

METHODOLOGY

In this section we explain each sub task involved in the pipeline and how we implemented each task within the final solution.

Vehicle Detection

In this stage, we analyse CCTV video footage one frame at a time. Each frame is considered as an independent image. Using computer vision and image processing techniques we can detect the object class we are interested in this work, vehicles. Vehicle detection is used in latter stages to estimate the change between two frames and eventually track the vehicle. In this phase, our main objective is to detect the vehicles which resides in a frame for a moment.

Prior to vehicle detection we performed background subtraction and blob detection with traditional image processing techniques from OpenCV and BGS Library. We experimented with detection techniques available in OpenCV library such as SIFT (Scale Invariant Feature Transform) [8], SURF (Speed Up Robust Feature) [9], ORB (Oriented Brief and Fast Key Descriptor) and Haar classifier [10].

Vehicle Tracking

Objective of this stage is to uniquely identify each vehicle through multiple frames. As given above, the output of vehicle detection becomes an input to this stage. Identifying each vehicle uniquely or tracking is important to estimate the traffic flow and also to detect certain traffic violation types.

To implement this stage we used widely used object tracking algorithms such as CAM-Shift (Continuously Adaptive Mean Shift) [11], Kalman Tracking and also algorithms provided in OpenCV such as BOOSTING, TLD (Tracking Learning and Detection), KCF (Kernelized Correlation Filter), GOTURN (Generic Object Tracking Using Regression), MEDIANFLOW and MIL (Multiple Instance Learning). With the best vehicle detection methodology we have integrated these vehicles tracking methods uniquely and independently to proceed with the best detection and tracking to cooperate with the classification process.

Vehicle Recognition and Counting

In this stage, the system uses to mark the nearest point or nearest distance of camera so it can detect the vehicles which are being reached at a frame. The process of vehicle detection begins while we process the image of the nearest vehicle. In this stage, we extract some specific features from the image and create a vector to represent this object we detected from the above step.

Once the processing of single frames which are considered to as images and extraction is being carried out, we can represent this single image as a feature vector. This feature vector uses to include all the features which are being extracted in the processing stage. This feature vector creation is handled through specific machine learning and deep learning methodologies which will address the concepts of classification. Once the classification is being carried out, with the summarized results we can present the categorical count of vehicles to the user.

The template is designed using after considering the deep learning models. We have use the inception V3 model which incorporates the pre-defined weights, max pooling modules performing dimensionality reduction with increase number of computations. We have trained the model to classify whether the object is a vehicle or not and to classify according to six predefined categories: bike, bus, car, van, trisho and lorry.

EVALUATION

In this section, we perform define the evaluation criteria that is being used for the series of experiments that have been conducted to validate the design and implementation of this project. We thought of calculating the accuracy of detection and tracking mechanisms with the classification test accuracy. Detection accuracy had been considered with the ground truth collection of the data and constituting the confusion matrix. When considering the accuracy measure for the visual tracking of multiple objects can be considered as a n active research field with many application domains. Recently there were a huge growth in performing the evaluations of tracking approaches. We can list down them as CHIL [15], AMI [12], U.S. VACE [13], ETISEO [14], U.K. Home office iLIDS [16], CAVIAR [17], CREDS [18], PETS [19], INCH [20] and CLEAR [21]. Due to high number of different metrics presented evaluation of object detection, localization and the tracking will have its own dependencies between separate metrics. In here there was issue of commonly agreed metric which also can generally applicable rather than using a custom set of measures to each metric. To solve this MOTP (Multiple Object Tracking Precision) and MOTA [22] (Multiple Object Tracking Accuracy) are being introduced to detect basic types of errors while expressing the tracker's overall strengths and the suitability of using the in the typical performance evaluation. This MOTA (Multiple Object Tracking Accuracy) defines the overall accuracy when considered with the ground truth values. When calculating the overall accuracy this metric use to consider the misses, mismatches and the false positives with the tracked objects. In here we could avoid considering the MOTP (Multiple Object Tracking Precision) as it defines the localization of the tracking procedure.

RESULTS

The project consists with three major phases. When considering the Vehicle Detection phase, we tried out 5 main methods. These methods are blob detection, SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features), FLANN (Fast Library for Approximate Nearest Neighbors), ORB (Oriented Fast Brief Key Descriptor) and Haar Classifier. Out of these methods blob detections produce an accuracy of 55.1% and Haar Classifier produce an accuracy of 72.79%. These two accuracies are calculated using a confusion matrix considering the false matches, misses and true matches while other methods denoted the number of key points detected when considered with the feeded image. SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features), FLANN (Fast Library for Approximate Nearest Neighbors) and ORB (Oriented Fast Brief Key Descriptor) produce 2, 12, 2 and 24 key point detections respectively. In the Tracking phase, we have tried using 7 methodologies. They are Improved version of CAM-Shift (Continuously Adaptive Mean Shift) which includes the multi detection, MIL (Multiple Instance Learning), TLD (Tracking Learning and Detection), KCF (Kernelized Correlation Filter), Medianflow, Boosting and Kalman Tracking. These methods produce the accuracy of 66.67%, 67.78%, 70%, 67.78%, 67.78%, 68.89% and 98.89% respectively. These tracking methods were evaluated against the ground truth and the resultant results from these algorithms using the Multiple Object Tracking Accuracy Algorithm.

In the classification phase, we have tried 2 different methods. They are using a positive/negative dataset and using a classified dataset including the basic vehicle types. These methods result 97.9% and 87.2% respectively using the test accuracies. Then we integrated the deep learning model into the haar classifier which results 77.78% of accuracy calculated using a confusion matrix.

When considering the haar classifier and the Kalman tracking pipeline the entire tracking accuracy was 75.56% calculated against the MOTA (Multiple Object Tracking Accuracy) algorithm.

CONCLUSION

In this paper, we have discussed the implementation and evaluation of a automated traffic monitoring system. When considering the vehicle detection results based on the above methods we have tested. When considering SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features), FLANN (Fast Library for Approximate Nearest Neighbors) and ORB (Oriented Fast Brief Key Descriptor) are well defined computer vision based feature descriptors. These methods do not imply a thresholding which defines an accuracy which we could convert into a percentage. When considering those methods, we could only say that when describing the detected features and matching those features ORB uses to be in an appropriate position based on the results.

When considering Blob Detection and Haar Classifier the accuracy was measured for a 60 seconds CCTV footage clip. As we can see using computer vision based methods haar classifier provides a better accuracy. When considering the tracking mechanisms. So, in here there were few methods like CAM-Shift (Continuously Adaptive Mean Shift), OpenCV methods and Kalman Tracking. When talking about CAMShift tracking we did number of improvements to use that for multi tracking purpose with unique ID generation. When talking about OpenCV methods they are fixed. Build with multi tracking purposes when the Region of Interest is given in the first frame. Once the Region of Interest moves away from the frame the bounding box use to get moved to the corner and sits there for other frames where those regions are not there. In here we cannot provide the regions as the source is not open for the users.

FUTURE WORK

The entire pipeline accuracy can be improved by utilizing a state-of-the-art object detection technique based on deep learning such as Faster-RCNN (Region Based Convolution Neural Network). This can be involving in higher level of feature matching when considering the pixel information, so the detection of each vehicle type would be in an equal high accuracy position.

REFERENCES

[1] "CCTV Division–Sri Lanka Police" [Online]. Available: https://www.police.lk/index.php/cctv-division. [Accessed August 2017]

[2] G. D. Sullivan, K. D. Baker, A. D. Worrall, C. I. Attwood and P. R. Remagnino, "Model-based vehicle detection and classification using orthographic approximations".

[3] F. M. d. S. Matos and R. M. C. R. d. Souza, "Hierarchical classification of vehicle images using nn with conditional adaptive distance," Neural Information Processing, ser. Lecture Notes in Computer Science, vol. 8227, p. pp. 745–752., 2013.

[4] J. Zheng, Y. Wang and M. E. Hallenbeck, "Extracting Roadway Background Image: a Mode-Based Approach," TRB 2006 Annual Meeting, 2006.

[5] V. R. S, B. L. N. P and A. K. R, "Computer Vision based Vehicle Recognition on Indian Road," International Journal of Computer Vision and Signal Processing, vol. 5, no. 1, 2015.

[6] Z. Sun, G. Bebis and R. Miller, "ON-ROAD VEHICLE DETECTION USING GABOR FILTERS AND SUPPORT VECTOR MACHINES".

[7] M. Kafai and B. Bhanu, "Dynamic Bayesian Networks for Vehicle Classification in Video," IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, vol. 8, no. 1, 2012.

[8] Y. Tian, Y. Wang, Z. Hu and T. Huang, "Selective eigenbackground for background modeling and subtraction in crowded scenes," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 11, pp. 1849-1864, 2013.

[9] L. Chen, J. Hsieh, H. Chiang and T. Tsai, "Real-time vehicle color identification using symmetrical SURFs and chromatic strength," International Symposium on Circuits and Systems, pp. 2804-2807, 2015.

[10] S. Elkerdawi, R. Sayed and M. ElHelw, "Real-time vehicle detection and tracking using Haar-like features and compressive tracking," 1st Iberian Robotics Conference, pp. 381-390, 2014.

[11] J. Xia, J. Wu, H. Zhai and Z. Cui, "Moving Vehicle Tracking Based on Double Difference and CAMShif," CAMShift", International Symposium on Information Processing (ISIP'09), pp. pp. 029-032, August 2009.

[12] "CHIL—Computers in the Human Interaction Loop," [Online]. Available: http://chil.server.de/. [Accessed August 2017]

[13] "AMI—Augmented Multiparty Interaction," [Online]. Available: http://www.amiproject.org/. [Accessed September 2017].

[14] "VACE—Video Analysis and Content Extraction," [Online]. Available: http://www.informedia.cs.cmu.edu/arda/vaceII.html. [Accessed August 2017].

[15] R. R. and J. M., "Real time classification and tracking of multiple vehicles in highways," Pattern Recognition Letters, pp. 1597-1607, 2005.

[16] "ETISEO—Video Understanding Evaluation," [Online]. Available: http://www.silogic.fr/etiseo/. [Accessed September 2017].

[17] "CAVIAR—Context Aware Vision using Image-based Active Recognition," [Online]. Available: http://homepages.inf.ed.ac.uk/rbf/CAVIAR/. [Accessed September 2017].

[18] F. Ziliani, S. Velastin and e. a. F. Porikli, "Performance evaluation of event detection solutions: the CREDS experience," Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS ’05), p. 201–206, 2005.
[19] "PETS—Performance Evaluation of Tracking and Surveillance," [Online]. Available: http://www.cbsr.ia.ac.cn/conferences/VS-PETS- 2005/. [Accessed August 2017].
[20] "EEMCV—Empirical Evaluation Methods in Computer Vision," [Online]. Available: http://www.cs.colostate.edu/eemcv2005/. [Accessed Octorber 2017].
[21] "CLEAR—Classification of Events, Activities and Relationships," [Online]. Available: http://www.clear-evaluation.org/. [Accessed November 2017].
[22] K. Nernardin and R. Stlefelghagen, "Evaluating Multiple Object Tracking Performance, The CLEAR MOT Metric," EURASIP Journal on Image and Video Processing, 2008.