DETECTION OF MOTORCYCLISTS WITHOUT HELMET IN IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

[1] Manyala prathyusha
Btech (CSE)

[2]G. parameshwari
Btech (CSE)

[3]Ch.sukumar
Btech (CSE)

[4]K.murali M.tech(PHD)
Associate professor
CMR TECHNICAL CAMPUS

ABSTRACT In order to ensure the safety measures, the detection of traffic rule violators is a highly desirable but challenging task due to various difficulties such as occlusion, illumination, poor quality of surveillance images, varying whether conditions, etc. In this paper, we present a framework for automatic detection of motorcyclists driving without helmets in surveillance images. In the proposed approach, first we use adaptive background subtraction on images frames to get moving objects. Later convolutional neural network (CNN) is used to select motorcyclists among the moving objects. Again, we apply CNN on upper one fourth part for further recognition of motorcyclists driving without a helmet. The performance of the proposed approach is evaluated on two datasets, IIT H Helmet 1 contains sparse traffic and IIT H Helmet 2 contains dense traffic, respectively.

INTRODUCTION Since, motorcycles are affordable and a daily mode of transport, there has been a rapid increase in motorcycle accidents due to the fact that most of the motorcyclists do not wear a helmet which makes it an ever-present danger every day to travel by motorcycle [1], [2]. In the last couple of years alone most of the deaths in accidents are due to damage in the head [3]. Because of this wearing helmet is mandatory as per traffic rules, violation of which attract hefty fines. Inspite, a large number of motorcyclists do not obey the rule. Presently, all major cities already deployed large video surveillance network to keep a vigil on a wide variety of threats. Thus using such already existing system will be a cost efficient solution, however these systems involve a large number of humans whose performance is not sustainable for long periods of time. Recent studies have shown that human surveillance proves ineffective, as the duration of monitoring of videos increases, the errors made by humans also increases. [4]. To date several researchers [5], [6], [7], [8], [9], [1], [2] have tried to tackle the problem of detection of motorcyclists without helmet by using different methods but have not been able to accurately identify motorcyclists without helmets under challenging conditions such as occlusion, illumination, poor quality of video, varying weather conditions, etc. One major reason of the poor performance of existing methods is the use of less discriminative representation for object classification as well as the consideration off irrelevant objects against the objective of detection of motorcyclists without helmet. Also, the existing approaches make use of handcrafted features only. Deep networks have gained much attention with state-ofthe-art results in complicated tasks such as image classification [10], object recognition [11], tracking [12], [13], detection and segmentation [14] due to their ability to learn features directly from raw data without resorting to manual tweaking. However, deep networks have not been explored till date for this task as per the best knowledge of the authors. The overall contribution of this paper is as follows: • Use of adaptive background modeling for the detection of moving vehicles on busy roads which handle the challenges such as illumination effects, weather change, etc.

• Instead of using hand-crafted features, we have explored the ability of convolutional neural network (CNN) to improve the classification performance.

•The proposed approach is evaluated on sparse traffic videos as used in [1], [2] as well as on crowded traffic videos collected from the CCTV Surveillance Network of the Hyderabad City, India edge histograms used circular hough transforms to compare and classify helmets, it leads to a lot of mis-classification among motorcyclists with helmet as helmet like objects were also classified as helmet as well as the helmets which were different were not classified as helmets. To overcome this mis-classification problem, Silva et al. [7], [9] proposed a system in which he tracks the vehicles using Kalman filter [15]. An important advantage of this Kalman tracking system [15] is the ability to continue to track objects even if they are lightly occluded but when there were more than two or three motorcyclists appear in a same frame, Kalman filter [15] fails because Kalman filter [15] mostly works well for linear state transitions (i.e tracking single objects/one object at a time). But to track multiple objects, we need non-linear functions to track them. Recently, Dahiya et al. [1] proposed a system which first uses Gaussian mixture model to detect moving objects. This model is robust to slight variations in the background. It uses two classifier in serial, one for the separating motorcyclist from moving objects and another for separating without helmet from the upper one fourth part of the motorcyclists. However, it uses only hand engineered features such as SIFT [16], HOG [17], LBP [18] along with kernel SVM in both classifications. Their approach was promising as it had accurately classified motorcyclists and non-motorcyclists but was not able to accurately classify between helmet and non-helmet riders under difficult conditions. Singh et al. [2] proposed a visual big data framework which scales the method in [1] to a city scale surveillance network. Experimental results shows that the framework is able to detect a violator in less than 10 milliseconds. The existing methods suffer from several challenges such as occlusion of objects and illumination effects as well as they tried to address it by using SVM [19], [20], [21] for classification between motorcyclists and non motorcyclists and helmet riders and without helmet riders which made localization of occluded objects easier. But for that to efficiently work, we also need to have good features from the motorcyclists to classify accurately which is difficult using HOG [17] or LBP [18] or SIFT [16] on images with less pixels. This inspired us to come up with a method, which uses CNN [22] to extract discriminative features.

PROPOSED FRAMEWORK FOR HELMET DETECTION

In the proposed system, first we apply adaptive background subtraction to detect the moving objects. These moving objects are then given to a CNN [22] classifier as input which then classifies them into two classes, namely, motorcyclists and nonmotorcyclists. After this, objects other than motorcyclists are discarded and passed only objects predicted as motorcyclist for next step where we determine weather the motorcyclist is wearing a helmet or not again using another CNN classifier. We assume that the head is located in the upper part of the incoming images and thus locate the head into top one fourth part of images. The located head of the motorcyclist is then given as input to second CNN which is trained to classify withhelmet vs. without-helmets. In the following subsections, we explain each step in details.

Convolutional Neural Network for Object Classification A convolutional neural network (CNN) is a variant of feed forward neural networks using back propagation algorithm. It learns high-level features from the spatial data like image. The recent widespread success of convolutional neural networks is in it's ability to extract inter-dependant information from the images i.e localization of the pixels which are highly sensitive to other pixels. The convolutional neural network training consist of convolution layers, relu layers maxpooling layers, fully connected layers and a loss function (e.g. SVM/Softmax) on the last (fully-connected) layer. In the primary layers we get the edge information of the images similar to some of the handcrafted algorithms but, In the final layers, we start getting texture and ridge information which helps us in getting sensitive information usefull for classification

Recognition of Motorcyclists without Helmet To recognize motorcyclists without helmet, from the images of motorcyclists, we cropped only the top one fourth part of the image as that was the region where the motorcyclist's head is located most of the time. From this, we locate the portion of the head by subtracting the binary image of the foreground of same region. Then we build a CNN model in order to separate the without-helmet from the with-helmet images. This model is trained for the binary classification of helmet and head. the feature maps of the sample helmets. These feature maps illustrate that the CNN learns the common hidden structures among the helmets in the training set and thus able to distinguished between a helmet and a head.

CONCLUSION

The proposed framework for automatic detection of motorcyclists driving without helmets makes use of adaptive background subtraction which is invariant to various challenges such as illumination, poor quality of video, etc. The use of the deep learning for automatic learning of discriminative representations for classification tasks improves the detection rate and reduces the false alarms resulting into more reliable system. The experiments on real videos successfully detect ≈ 92.87% violators with a low false alarm rate of ≈ 0.50% on two real video datasets and thus shows the efficiency of the proposed approach.

REFERENCES

[1] K. Dahiya, D. Singh, and C. K. Mohan, "Automatic detection of bikeriders without helmet using surveillance videos in real-time," in Proc. Int. Joint Conf. Neural Networks (IJCNN), Vancouver, Canada, 24–29 July 2016, pp. 3046–3051.

[2] D. Singh, C. Vishnu, and C. K. Mohan, “Visual big data analytics for traffic monitoring in smart city,” in Proc. IEEE Conf. Machine Learning and Application (ICMLA), Anaheim, California, 18–20 December 2016.

[3] C. Behera, R. Ravi, L. Sanjeev, and D. T, “A comprehensive study of motorcycle fatalities in south delhi,” Journal of Indian Academy of Forensic Medicine, vol. 31, no. 1, pp. 6–10, 2009.

[4] W. Hu, T. Tan, L. Wang, and S. Maybank, “A survey on visual surveillance of object motion and behaviors,” IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 34, no. 3, pp. 334–352, 2004.

[5] C.-C. Chiu, M.-Y. Ku, and H.-T. Chen, “Motorcycle detection and tracking system with occlusion segmentation,” in Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services, Santorini, Greece, 6–8 June 2007, pp. 32–32.

[6] J. Chiverton, “Helmet presence classification with motorcycle detection and tracking,” IET Intelligent Transport Systems (ITS), vol. 6, no. 3, pp. 259–269, 2012.

[7] R. Silva, K. Aires, T. Santos, K. Abdala, R. Veras, and A. Soares, “Automatic detection of motorcyclists without helmet,” in Proc. Latin American Computing Conf. (CLEI), Puerto Azul, Venezuela, 4–6 October 2013, pp. 1–7.

[8] W. Rattapoom, B. Nannaphat, T. Vasan, T. Chainarong, and P. Pattanawadee, “Machine vision techniques for motorcycle safety helmet detection,” in Proc. Int. Conf. Image and Vision Computing New Zealand (IVCNZ), Wellington, New Zealand, 27–29 November 2013, pp. 35–40.

[9] R. V. Silva, T. Aires, and V. Rodrigo, “Helmet detection on motorcyclists using image descriptors and classifiers,” in Proc. Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, 27–30 August 2014, pp. 141–148.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proc. Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, Nevada, United States, 3–6 December 2012, pp. 1097–1105. [11] D. Jeff, J. Yangqing, V. Oriol, H. Judy, Z. Ning, T. Eric, and D. Trevor, “DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition,” Int. Conf. on Machine Learning (ICML), vol. 32, no. 1, pp. 647–655, 2014.

[12] N. Hyeonseob and H. Bohyung, “Learning Multi-Domain Convolutional Neural Networks for Visual Tracking,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, United States, June 26th - July 1st 2016, pp. 4293–4302.

[13] Z. Kaihua, L. Qingshan, Wu, and Y. Ming-Hsuan, “Robust Visual Tracking via Convolutional Networks without Training,” IEEE Trans. Image Processing, vol. 25, no. 4, pp. 1779–1792, 2016.

[14] G. Ross, D. Jeff, D. Trevor, and M. Jitendra, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Colombus, Ohio, 24–27 June 2014, pp. 580–587.

[15] R. E. Kalman, “A new approach to linear filtering and prediction problems,” Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45, 1960.

[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[17] D. Navneet and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), San Diego, California, 20–26 June 2005, pp. 886– 893.

[18] Z. Guo, D. Zhang, and L. Zhang, "A completed modeling of local binary pattern operator for texture classification," IEEE Trans. Image Processing, vol. 19, no. 6, pp. 1657–1663, 2010.

[19] C. Cortes and V. Vapnik, "Support vector networks," Machine Learning (Springer), vol. 20, no. 3, pp. 273–297, 1993.

[20] D. Singh, D. Roy, and C. K. Mohan, "Dip-svm:distribution preserving kernel support vector machine for big data," IEEE Trans. on Big Data, 2017. [Online]. Available: http://dx.doi.org/10.1109/TBDATA.2016.2646700