

Detection of Motorcyclists without Helmet in Videos using Convolutional Neural Network

C. Vishnu, Dinesh Singh, C. Krishna Mohan and Sobhan Babu

Visual Intelligence and Learning Group (VIGIL), Department of Computer Science and Engineering
Indian Institute of Technology Hyderabad, Kandi, Sangareddy-502285, India

Email: {cs16mtech11021, cs14resch11003, ckm, sobhan}@iith.ac.in

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 video, varying whether conditions, etc. In this paper, we present a framework for automatic detection of motorcyclists driving without helmets in surveillance videos. In the proposed approach, first we use adaptive background subtraction on video 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, *IITH_Helmet_1* contains sparse traffic and *IITH_Helmet_2* contains dense traffic, respectively. The experiments on real videos successfully detect 92.87% violators with a low false alarm rate of 0.5% on an average and thus shows the efficacy of the proposed approach.

Keywords—*Helmet Detection, Traffic Surveillance, Deep Learning, Convolutional Neural Network.*

I. 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. In spite, 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-of-the-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.

The remainder of this paper is organized as follows. Section II presents the related work. Section III describes proposed approach for automatic detection of motorcyclists without helmets. Section IV discusses the experimental setup, dataset, and performance. Finally, we conclude in section V.

II. RELATED WORK

To date many researchers have proposed several methods [5], [6], [7], [9], [8], [1], [2] to solve this problem of real time helmet detection in traffic. These methods are discussed below in this section.

Chiu *et al.* [5] proposed a system to solve the motorcyclists detection in surveillance videos. This system segments the moving object and then tracks motorcycles and heads using a probability-based algorithm which handle the occlusion problem but unable to handle small variations due to noise and illumination effects. Also, it uses Canny edge detection with a search window of certain size in order to detect head. Chiverton *et al.* [6] used edge histogram based features in order to detect motorcyclists. The strength of this method is that it performs well even if there was low light or low illumination in videos due to the use of edge histograms near the head instead of detecting the features of the head region. Since the

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.

III. PROPOSED FRAMEWORK FOR HELMET DETECTION

Fig. 1 shows the block diagram of the proposed system. 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 *non-motorcyclists*. 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 *with-helmet* vs. *without-helmets*. In the following subsections, we explain each step in details.



Fig. 1. Block diagram of proposed framework for the detection of motorcyclists without Helmet

A. Background Modeling and Moving Object Detection

First, we apply background subtraction method to separate moving objects such as motorcycle, humans, cars from traffic videos using improved adaptive Gaussian mixture model in [23] which is robust to certain challenges like illumination variance over the day, shadows, shaking tree branches and other sudden changes. We use variable number of Gaussian models for each pixel because single Gaussian is not sufficient to completely model these variations in complex and variable situations [24]. Here we provide a brief overview of the improved adaptive Gaussian mixture model.

Let us consider I^1, I^2, \dots, I^t be the intensity of a pixel for past t consecutive frames. Then at time t , the probability of observing intensity value for a pixel is given by:

$$P(I^t) = \sum_{j=1}^K w_j^t \times \eta(I^t, \mu_j^t, \sigma_j^t), \quad (1)$$

where, w_j^t is weight and $\eta(\cdot, \cdot, \cdot)$ is j^{th} Gaussian probability density function with mean μ_j^t and σ_j^t as variance at time t . For each pixel, the Gaussian components with low variance and high weight correspond to background class and others with high variance correspond to foreground class. At time t , the pixel intensity I^t is checked against all Gaussian components. If j^{th} component satisfies the condition :

$$|\mu_j^t - I^t| < e_j \sigma_j^t, \quad (2)$$

then j^{th} component is considered to be a match. Also, the current pixel is classified as background or foreground according to the class of j^{th} Gaussian model. The weight update rule is given by :

$$w_j^t = (1 - \alpha)w_j^{t-1} + \alpha(M_j^t), \quad (3)$$

$$M_j^t = \begin{cases} 0, & \text{for matched model} \\ 1, & \text{otherwise,} \end{cases} \quad (4)$$

where, α is learning rate which determines how frequently parameters are adjusted. Here, e_j is a threshold which has significant impact when different regions have different lighting. Generally, the value of e_j is kept around 3, as $\mu^t \pm 3\sigma_j^t$ accounts for approximately 99% of data [23]. Also, other parameters of matched models are updated as:

$$\mu^t = (1 - \rho)\mu^{t-1} + \rho I^t, \quad (5)$$

$$(\sigma^2)^{(t)} = (1 - \rho)(\sigma^2)^{(t-1)} + \rho(I^t - \mu^t)^2. \quad (6)$$

Here, $\rho = \eta(I^t | \mu_j, \sigma_j)$. When there is no matched component, a new Gaussian model is created with current pixel value as mean, low prior weight and high variance. This newly created model replaces the least probable component or added as a new component if maximum number of components is reached or not, respectively. All the moving objects (i.e. foreground objects) are resized to a fixed size before giving them as input to a CNN classifier.

B. 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 its 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.

C. Recognition of Motorcyclists from Moving Objects

To find bounding boxes of different objects, we used Gaussian background subtraction which uses a method to model each background pixel by a mixture of K Gaussian distributions ($K = 3$ to 5). The probable background colours are the ones which stay longer and are more static. On these varying pixels, we draw a rectangular bounding box. After obtaining all the objects of motorcyclists and non-motorcyclists, a CNN model is built using these images to separate the motorcyclists from other moving objects. Fig. 2 show the feature maps of the sample motorcycles. These feature maps illustrate that the CNN learns the common hidden structures among the motorcyclist in the training set and thus able to distinguish between a motorcyclist and other objects.

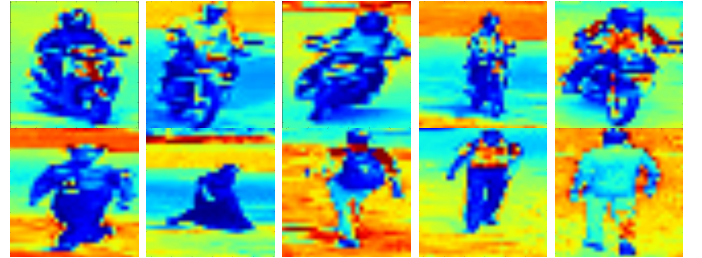


Fig. 2. Visualization of the trained representation by CNN for the classification of motorcycle and not-motorcycle

D. 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. Fig. 3 shows 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.

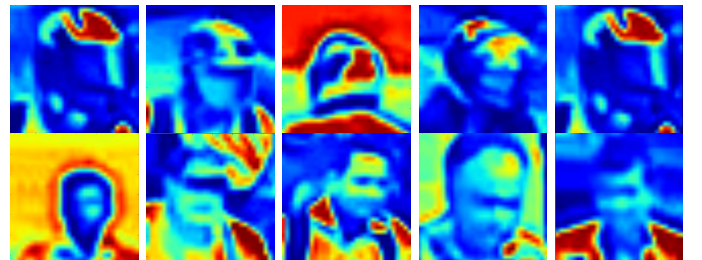


Fig. 3. Visualization of the trained representation by CNN for the classification of with-helmet and without-helmet

IV. EXPERIMENTAL EVALUATION

The experiments are conducted on a machine running Ubuntu 16.04 Xenial Xerus having specifications Intel(R) Xeon(R) CPU E5-2697 v2 @ 2.70GHz×48 processor, 128GB RAM with NVIDIA Corporation GK110GL [Tesla K20c]×2

GPUs. The programs for helmet detection are written in *Python* – 2.7.12 with the help of the various libraries such as *OpenCV* – 3.0. for image processing and vision tasks, *Keras* – 1.1.1 [25] a deep learning library to train CNN models, *Theano*–0.8.2, *Scikitlearn*–0.18, *Numpy*–1.11.2 for maths and linear algebra operations. Here, the value of K the number of Gaussian components for each pixel is kept in between 3 and 5, which is determined empirically. All moving objects are resized to 32×32 before giving them input to CNN. The architecture for the CNN is same as used in [10] for CIFAR dataset.

A. Datasets Used

The performance of the proposed approach is evaluated on two video datasets containing sparse traffic and dense traffic, respectively.

IITH_Helmet_1: This dataset is collected from the surveillance network at Indian Institute of Technology Hyderabad, India (IITH) campus because there was no public data set available at the start of this research work. It is a 2 hour surveillance video data which is collected at 30 frames per second. Fig. 4 presents sample frames from the collected dataset. We have used the first one hour of the video for



Fig. 4. *IITH_Helmet_1*: Video dataset for helmet detection in sparse traffic collected from the CCTV surveillance network of IIT Hyderabad campus [1], [2].

training the model and the remaining for testing purpose. The training video contains 42 motorcycles, 13 cars, and 40 humans. Whereas, testing video contains 63 motorcycles, 25 cars, and 66 humans. Fig. 5 shows the 2D visualization of spread of the extracted features for Motorcyclists vs. Non-motorcyclists using t-SNE [26]. Fig. 6 shows the 2D visualization of spread of the extracted features for ‘Helmet’ vs. ‘Non-Helmet’. Here, the classification of the motorcyclists vs other objects is relatively easy to classify because the patterns corresponds to other objects are deviating significantly from the patterns of motorcyclists. But the deviation is very less among the patterns corresponding to head and helmets (i.e. two classes are overlapping) which make the classification task more complex.

IITH_Helmet_2: This second dataset is acquired from the CCTV surveillance network of Hyderabad city in India. It is a 1.5 hour video which is collected at 25 frames per

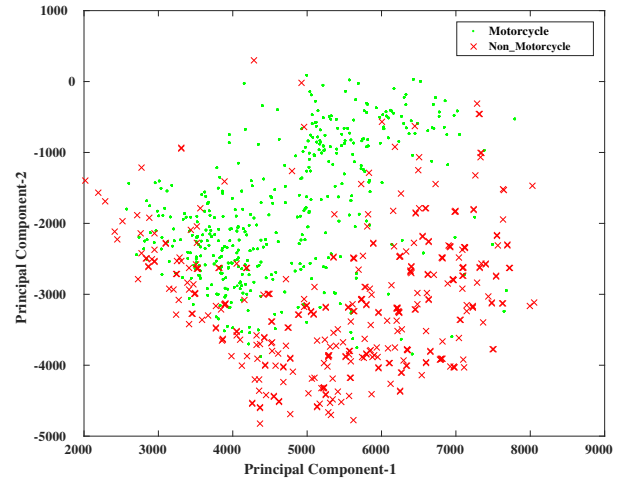


Fig. 5. 2D visualization of spread of the extracted features for ‘Motorcyclists’ vs. ‘Non-motorcyclists’ using t-SNE [26]

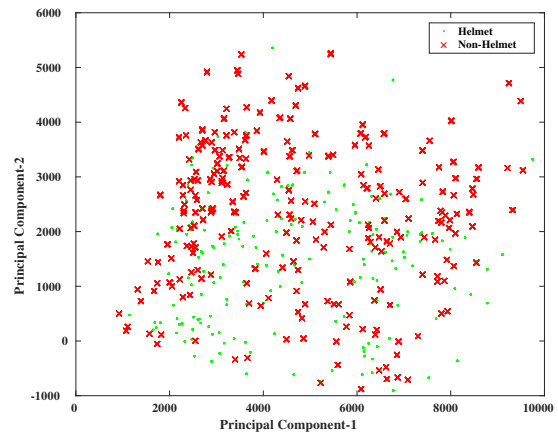


Fig. 6. 2D visualization of spread of the extracted features for ‘Helmet’ vs. ‘Non-Helmet’ using t-SNE [26]

second. The sample frames from this dataset are presented in Fig. 7. The first half an hour of the video is used for training the model and the remaining for testing purpose. The training video contains 1261 motorcyclists and 4960 non-motorcyclists. Whereas, testing video contains 2312 motorcyclists, and 9112 non-motorcyclists. Fig. 8 shows the 2D visualization of spread of the extracted features for Motorcyclists vs. Non-motorcyclists. Fig. 9 shows the 2D visualization of spread of the extracted features for ‘Helmet’ vs. ‘Non-Helmet’. Here, the classification of the motorcyclists vs other objects is easy to classify as most of the patterns corresponds to other objects are deviating significantly from the patterns of motorcyclists while few are very close thus they poses little challenge. While, the deviation is very less among the patterns corresponds to head and helmets (i.e. two classes are overlapping) which make the classification task more complex.



Fig. 7. *IITH_Helmet_2*: Video dataset for helmet detection in crowded traffic collected from the CCTV surveillance network of Hyderabad city in India.

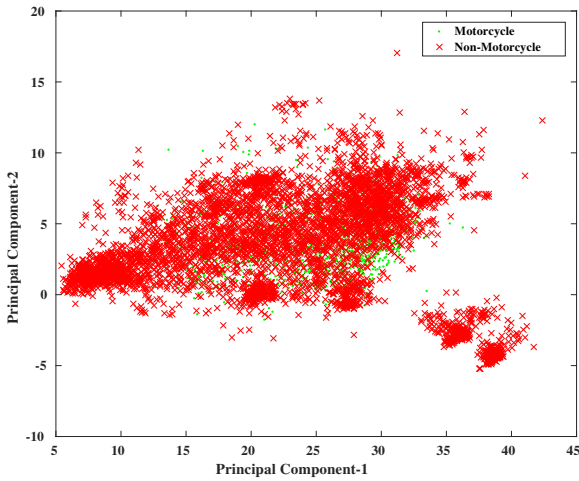


Fig. 8. 2D visualization of spread of the extracted features for ‘Motorcyclists’ vs. ‘Non-motorcyclists’ using t-SNE [26]

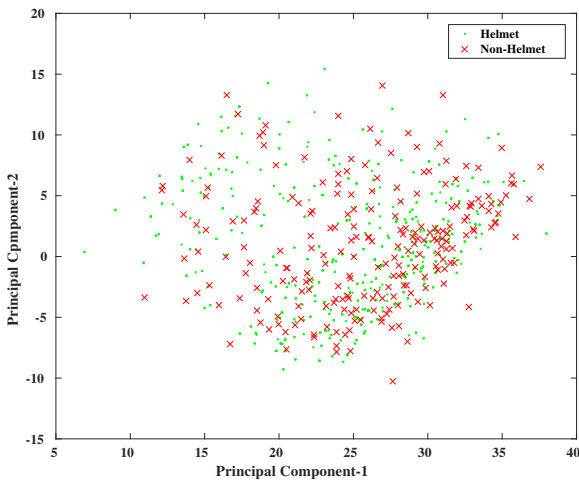


Fig. 9. 2D visualization of spread of the extracted features for ‘Helmet’ vs. ‘Non-Helmet’ using t-SNE [26]

B. Results and Discussion

In this section, we present experimental results and discuss the suitability of the best performing representation and model over the others. The architecture our model is based on AlexNet [10] consisting of 4 convolution layers with with 5 ReLU activation units, 2 max-pooling layers with dropout, and 2 fully-connected dense layers, with final *softmax* for classification into two classes.

The 5-fold cross validation is used to conduct experiments in order to have fair validation of the performance of the proposed approach. Table I presents the the results of the experiments for the classification of ‘Motorcyclist’ vs. ‘Non-motorcyclist’ using proposed CNN and the existing method used for comparison for both the datasets. For comparison we consider only HOG-SVM as its performance is highest among all other methods presented in [1]. Experiments show

TABLE I. PERFORMANCE (%) OF THE CLASSIFICATION OF ‘MOTORCYCLIST’ VS. ‘NON-MOTORCYCLIST’ USING CNN

DataSet:Feature	Fold1	Fold2	Fold3	Fold4	Fold5	Avg.(%)
<i>IITH_Helmet_1</i> :CNN	99.06	99.34	99.39	99.15	99.28	99.24
<i>IITH_Helmet_1</i> :HOG	97.93	99.59	98.35	99.38	99.17	98.88
<i>IITH_Helmet_2</i> :CNN	91.81	91.79	91.84	91.85	91.78	91.81
<i>IITH_Helmet_2</i> :HOG	81.83	81.58	81.97	81.23	82.59	81.84

that the accuracy is 99.24% with a low false alarm rate less than 0.5% on *IITH_Helmet_1* dataset and 91.81% with a low false alarm rate less than 0.5% on *IITH_Helmet_2* dataset. The proposed method using CNN outperforms the classification performance of the existing HOG-SVM with a margin of 0.36% on *IITH_Helmet_1* dataset and 9.97% on *IITH_Helmet_2* dataset, as illustrated in Fig. 10.

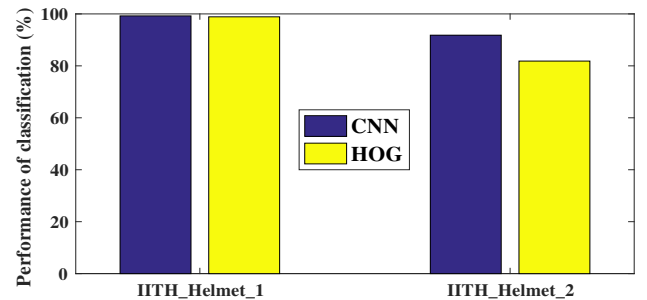


Fig. 10. Performance comparison of classification (%) of ‘motorcyclists’ vs. ‘non-motorcyclist’ in proposed approach using CNN with HoG-SVM [1].

For the second classification also we used 5-fold cross validation in order to validate the performance of proposed and existing methods. Table II lists the results of the experiments for the classification of ‘Helmet’ vs. ‘Non-Helmet’ using proposed CNN and the existing method used for comparison for both the datasets. For comparison we consider only HOG-SVM as its performance is highest among all methods presented in [1]. Experiments show that the accuracy is 98.63% with a low false alarm rate less than 0.5% on *IITH_Helmet_1* dataset and 87.11% with a low false alarm rate less than 0.5% on *IITH_Helmet_2* dataset. The proposed method using CNN outperforms the classification performance of the existing HOG-SVM with a margin of 4.83% on *IITH_Helmet_1* dataset and 29.33% on *IITH_Helmet_2* dataset, as illustrated in Fig. 11.

TABLE II. PERFORMANCE (%) OF THE CLASSIFICATION OF ‘HELMET’ VS ‘WITHOUT HELMET’ USING CNN

Dataset:Feature	Fold1	Fold2	Fold3	Fold4	Fold5	Avg.(%)
IITH_Helmet_1:CNN	98.73	98.65	98.61	98.48	98.69	98.63
IITH_Helmet_1:HOG	90.12	95.06	93.83	95.00	95.00	93.80
IITH_Helmet_2:CNN	87.28	86.85	87.32	86.95	87.18	87.11
IITH_Helmet_2:HOG	56.88	55.50	63.76	54.50	58.26	57.78

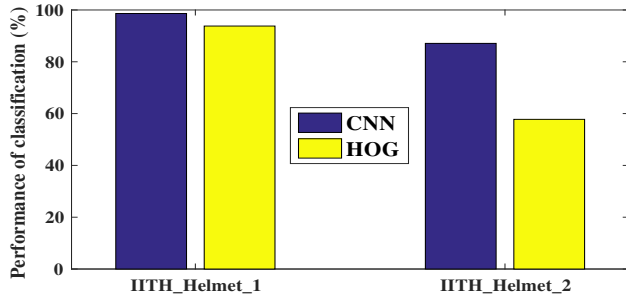


Fig. 11. Performance comparison of classification (%) of ‘Motorcyclist with helmet’ vs ‘Motorcyclist without helmet’ in proposed approach using CNN with HoG-SVM [1].

The final outcome of the experimental evaluation shows that using CNN improves the classification performance for both the classification tasks and thus leads to more reliable detection of violators driving without helmets. This major improvement is achieved for the classification of ‘Helmet’ Vs ‘Non-Helmet’.

V. 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 $\approx 92.87\%$ violators with a low false alarm rate of $\approx 0.50\%$ 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 bike-riders 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. Patanawadee, “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)*, Columbus, 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>
- [21] D. Singh and C. K. Mohan, “Distributed quadratic programming solver for kernel SVM using genetic algorithm,” in *Proc. IEEE Congress on Evolutionary Computation*, Vancouver, July 24–29 2016, pp. 152–159.
- [22] Y. Lecun, L. bottou, Y. bengio, and P. haffner, “Gradient-based learning applied to document recognition,” *Proceedings of IEEE*, vol. 86, no. 1, pp. 1–6, 1990.
- [23] Z. Zoran, “Improved adaptive gaussian mixture model for background subtraction,” in *Proc. Int. Conf. Pattern Recognition (ICPR)*, Cambridge, England, UK, 23–26 August 2004, pp. 28–31.
- [24] S. Chris and G. Eric, “Adaptive background mixture models for real-time tracking,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Collins, CO, USA, 23–25 June 1999, pp. 246–252.
- [25] F. Chollet, “Keras,” <https://github.com/fchollet/keras>, 2015.
- [26] V. der Maaten, Laurens, Hinton, and Geoffrey, “Visualizing data using t-sne,” *Journal of Machine Learning Research (JMLR)*, vol. 9, no. 1, pp. 2579–2605, 2008.