Proceedings of the International Conference on Smart Electronics and Communication (ICOSEC 2020)
IEEE Xplore Part Number: CFP20V90-ART; ISBN: 978-1-7281-5461-9

Real Time Automatic Detection of Motorcyclists With and Without a Safety Helmet

Valanukonda Lakshmi Padmini
Department Of Computer Science and Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada
valanukonda1999@gmail.com

Dr. G. Krishna Kishore
Associate Professor
Department of Computer Science and Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada,India
gkk@vrsiddhartha.ac.in

Ponnuru Durgamalleswarao
Department Of Computer Science and Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada
durga.durga.988722@gmail.com

Parasa Teja Sree
Department Of Computer Science and Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada
tejasreep89@gmail.com

***Abstract*: In the developing countries like India, the motorcycle riders are increasing day-by-day, wherein it also constitutes to the unprecedented increase in the number of motorcycle accidents across the country. To overcome this drawback, the proposed research work explains and demonstrates a method to enforce better safety protocols through the automatic detection of motorcyclists with and without a safety helmet by using a real-time traffic surveillance footage. The real-time automatic detection of motorcyclists with and without a safety helmet is established through detecting a vehicle and track pipelining it with OpenCV, sklearn, utilizing a descriptor known as the histogram of oriented gradients (HOG), and support vector classification (SVC), which are the combination of tools pertaining to machine learning and image processing mechanisms. With OpenCV Library method, a bike rider is identified in the surveillance video. Further by using a popular machine learning algorithm model called LinearSVC, the classifier label identifies whether the rider is wearing a safety helmet. The data attained in correspondence to the count of bike riders with and without safety helmet is stored in MySQL database with respective timestamps and is also visualized through tabular and graphical views in the developed desktop interface application. With 87.6% model accuracy, our paper proposes a solution to enhance the existing safety measures and provide a time-efficient approach to handle traffic regulations.**

Keywords: **Machine learning, image processing, histogram of oriented graphics (HOG), OpenCV, support vector classification, MySQL Connector, LinearSVC, NumPy, Pandas, sklearn**

I. INTRODUCTION

Across the world, Two-wheeler has become a popular passage. India with 1.3 billion population has millions of motor vehicles running on the roads. There is a high danger involved due to less protection. In the past 10 years, there was a rise in the count of accidents occurring due to motorcyclists without a helmet. Nowadays, accidents happening on the roads are one of the significant reasons for human deaths. Among various kinds of accidents that occur, Motorcycles accidents are abundant and cause serious injuries for both young and adult [7]. To reduce the risk, the motorcyclists should wear a helmet. The safety helmet is the motorcyclist's fundamental protection from major accidents and injuries. But a large number of the two-wheelers neglect the helmets during their rides. The Ministry of the country has made it blameworthy to ride a bike without a helmet and implementing many plans of action to snatch the lawbreakers. The traditional means [15] of checking that bikers with helmets are done by traffic police standing at a junction or by the captured CCTV flicks and then impose fines to those who violate the laws, which takes the human efforts and time-consuming. Automation of this process is highly required for the government to impose penalties on the violators, reduce work-force and time. The smart method one can actually think of is having a sensor engine that only starts when the helmet is properly worn by the person. But, the cost of this setup could be higher than most of the individuals of India could not effort it. To effectively address all these problems, we would like to bring an efficient approach for automatic helmet detection using Machine learning and image processing techniques. This method is considered to be efficient as there are CCTV's installed on roads and consume less time and reduce the need for large human force.

The suggested approach focusses on the absolute protection to the motorcyclists. Though helmet wearing is mandatory, many of the two-wheeler riders violate it, which leads to increment in the deaths occur every year. Hence, automatic helmet detection is the liable solution for the government to lay hold of violators ensuring the public safety. In the idea to make traffic and accident system acute, where it provides a desktop application, which can reduce the efforts of traffic police and provide the functionality to them to identify the number of violators driving on the roads. The proposed application will automatically extract the images of motorcyclists who refused to wear helmets and our system will save the images for further procedure. Images will be extracted from the surveillance video [13] on public roads. Our application will also provide features like the visualization of data related to the occurrence of vehicles on different attributes. The count of two-wheeler vehicles with and without a safety helmet is stored in MySQL database with timestamps to accurately verify the law violators. The proposed application automatically captures the bikers without helmet using real-time video monitoring and helps the police department to impose a penalty on the motorcyclists, who violate the rules of wearing a helmet while driving on roads.

II. LITERATURE SURVEY

In [1], using Wide-angle and Zoom cameras, detection of the bikers with and without a helmet is depicted. The wide-angle lens captures the image intersection that detects the biker in a frame of width less than 80 and height less than 120 using the HAAR algorithm, which is the first step and zoomed lens to detect the license plates. The next step is to verify all the frames whether a biker is having helmet or not and creates the boundaries to the head of a biker of height half of the area and width one-third of the area of the frame, where the biker is detected. If the biker lacks the helmet, the zoom lens spots the number plate and stores the related data and pictures in the database, which acts as a proof for traffic wing to impose fines. The obtained output is compared with manual outcomes with 0.05 significance.

In [2], by using segmentation, classification of moving objects, and helmet detection, and a biker is spotted with and without safety helmet. Segmentation of moving objects requires a background picture, which is reformed by using adaptive mixture of Gaussians, diminishes the processing time and false-positives in classifiers. The obtained outcome is transformed to feature vector, which helps in classifying the vehicles like bikes and non-bikes by using descriptors. Finally, helmet identification is done with Region of Interest (ROI) determination to reduce computational capita and helmet identification, image classification using six different classifiers, and extraction of features using three types of descriptors. MLP using HOG obtained the best outcome with 0.9137 accuracy

In [3], the paper focuses on deep learning strategies to classify the bikers with and without a helmet. The evolution in CNN and SSD capable of segmentation and classification of images, once in runtime by creating bounding boxes. Classification of images with and without a helmet is done by CNN models like VGG19, VGG16, Mobile Nets, and Google Nets (Inception_v3). Neural Net is trained with python TensorFlow and accuracy is determined. SSD model can collaborate with other nets structures such as Google Net and Mobile Nets into one net which constructs structure faster. Among the four CNN models, Mobile Nets got the best outcome.

In [4], the fully enhanced real-time structure determines motorcyclists with and without helmets from vigilance using HOG as classifier and LBP as a feature vector. Initially, each frame is studied from footage and extraction of vehicles moving is obtained by eliminating unnecessary areas, then extracted pictures are enhanced by using an adaptive Gaussian mixture. After performing the required preprocessing steps, the classification is done to minimize the capita. Two-stage classification is done to acquire the best output and reduce misclassification. Finally, the biker's head is obtained by calculating the ROI of the upper twenty percent of the picture and converted them to grayscale pictures. Segmented pictures are then sent to CNN for the classification of bikers with and without the helmet.

In [5], using HOG as a feature vector and SVM as a classifier, the classification of bikers having and not having a helmet is done in real-time. Background modeling utilizing background subtraction is done at the beginning which contains an extraction of features and classification to differentiate moving objects like cars, humans, and bikes from non-moving like roads, trees, and builds which helps in minimizing the efforts in next phases. In the second step classification of the biker is done. Different varieties of features and kernels are combined to inspect classification outputs. An average of 11.58 milliseconds is taken to process a video of 30 frames per seconds

III. BASIC CONCEPTS

Classification is implemented to identify the set of classes that they are belongs to, by the data that is used for training or the instances of whose category or group is already known. There are two major categories in the classification. They are

Supervised classification is human-guided classification. The outcome is based on the data instance or the image that are representative of specific classes or groups. The training sets are created based on the user knowledge.

Unsupervised classification is a software-based classification. The outcome is based on the analysis of software without providing the classes by the user. The user will provide the number of output classes and the software that is used.

Support Vector Machine (SVM) [8] Classifier is a supervised classification algorithm. It is used for two-group classification problems. Support Vector Machine is a fast and dependable algorithm that provides accurate result when the input data is limited and sufficient. SVM uses hyperplanes to classify the classes of data points.

Hyperplanes are decision boundaries that helps to classify the classes of the datasets. Data that is falling on one phase or side of the hyperplane is considered as one class and other side is classified as another class.

The aspect of the hyperplane will be decided by the count of features. The hyperplane will be a straight line, if the input features are two. The hyperplane will be a 2-D plane, if the input features is three. If the features exceed three, then it will be difficult to use the hyperplanes.

The Object detection [9] is a technology that is used to detect the things or objects from the input data that are given to the system. Object detection has multiple applications like face detection, self-driving cars, obstacle detection, security systems etc.

Feature Extraction [11] will reduce the number of resources i.e. memory, computation power etc. that are required to describe the data while applying classification algorithm. Classification algorithm may overfit if the large number of variables are used in the classification.

The system requires large memory and computation power if we used the data directly without applying the pre-processing techniques on the input data. The accuracy of the model will depend on the methods that are used to decrease the number variable as possible.

Image resizing is the procedure to change the image from one size to another i.e. from m x n to a x b. The resizing can be done by scaling, cropping, padding.

In scaling, we use some factors called as scale factors to change the size of the image. In cropping, we cut the pixels based on the required size. In padding, we add the dummy pixels based on the required size of the image.

Image blurring technique is used to highlight the required portion of the image. If we observe the blurred image carefully, the image seems like smooth and the edges are not observed. Generally, low pass filter is used for the blurring. The low frequency will be considered but the high frequency will be prevented by this filter. Here frequency is the difference in pixel values.

A very important step required to recognize the region of interest is to extract the pixels that belongs to the region of interest.

Image thresholding is an era of image segmentation that separate the image by converting the grayscale image into the binary images [6]. It is effective for the images those have the high level of contrast.

IV. PROPOSED METHODOLOGY

The proposed system can be divided into the two main parts. The first part consists of segmentation and the second on consists of classification of images. The first step is used to capture the moving objects in the surveillance camera. The vehicles are classified into two classes because it is necessary to know that vehicle is motorcycle or not. And the second step consists of helmet detection [12]. For vehicle classification and helmet detection we used the Support Vector Classification (SVC). Figure 1 describes each and every step of the system to detect the motorcyclist's with and without helmet.

Proposed System:

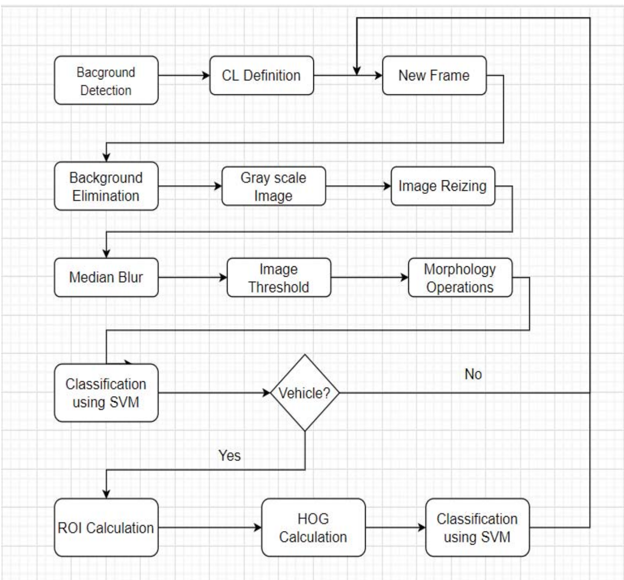

Figure 1: Proposed System

1. Background Detection:

To detect the moving objects, we have to detect the background of the scene. This background is used to extract the features from each frame. The frames are captured using the surveillance camera and these frames are used as a background. The static objects in the environment i.e. trees, parked vehicles, buildings etc. are captured as a background and used for further process. Fig 2.a indicates the image that contains the background.

2. Moving Object Segmentation:

This process reduces the processing time because it eliminates the unnecessary part of the frames and provides the outcome for the classification. In this step we will detect the interested area from the image and given as an input for further process. In this step we have to provide a cross line (CL), it is marked by us at a time we are starting the system. This cross line (CL) must be cross the road. When a vehicle crosses this CL, the moving object [14] segmentation process will start.

The frame is captured from the surveillance camera and the background elimination will be done by subtracting the captured frame with the background image. Fig 2.b and 2.c shows the captured frame and background eliminated image respectively. The background eliminated image is converted into the grayscale image. The Gray scale image should be resized to the required size using cropping or padding.

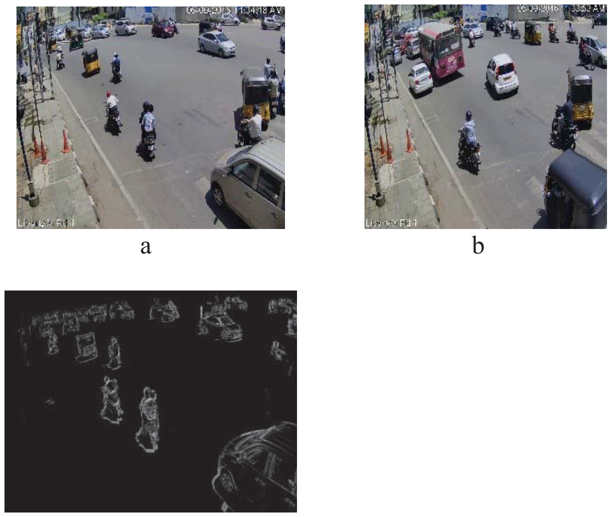

Figure 2: a) Background Scenario b) Frame c) Background Elimination

The median blur operation is applied to the resized image. This median blur will replace the center portion of the image by the median of all the pixels. The image threshold operation will be applied by the threshold value of 40. Lastly, the morphological operations i.e. Dilate and Erode are applied to remove the noises.

3. Vehicle Classification:

Before classifying the image, It is necessary to extract the features of the images. To extract the features from the frames, we use descriptors. The descriptors provides set of values as an outcome. This is called as "feature vector". This feature vector is used by the classifier to classify the each frame i.e. outcome of the previous steps.

Histogram Oriented Gradients (HOG) descriptor is used to extract the features from the images or frames.

The aim of any classifier is to use the features of the object to identify the object that what class it belogs to. To classify the vehicle we used the SVM Classifier[8]. This classification consists two classes of motorcycle and non-motorcycle. If the outcome of the classifier is 0, then the frame contains a vehicle but not a motorcycle. If the outcome of the classifier is 1, then there is a motorcycle in the frame. So it goes to the helmet detection phase. If the outcome is 2, then the frame doesn't contain a vehicle.

4. Helmet Detection:
Before detecting the helmet, we have to calculate the Regoin of Interest (RoI), Feature Extraction, and Image classification. The Region of Interest (RoI) is an important part in helmet detection. The motorcyclist's head must be in the part of RoI. Because of RoI the computation cost and time are drastically decreased. Actually top of the image is used as a region of interest.
For extracting the features from the frames, we used the descriptor named HOG. The gray scale image was used for the feature extraction process. HOG descriptor has bees set up with a 16 histograms by 16 partition window. This way, a vector of 256 features are generated. Here, we can use large variation of histogram and partiton window sizes to get the best results.
For image classification process[10], we used Linear SVM classifier. If the outcome of the helmet detection classifier is 0, then the motorcyclist with a safety helmet. Otherwise the motorcyclist without a safety helmet..

V. Implementation and Results

Step 1:

a. Machine Learning algorithms use data as their source of learning. To attain a well-detailed set of data points, we have based the CCTV surveillance footage as our foundation to develop the algorithm, since the road user summative and swift motorization are similar to the situation of the real-world.

b. The recorded samples of our data are huge in number, as shown in table 1. To train a model of such high complexity is resource-wise impossible. To alter the situation, we have extracted 1000 video clip samples to form the necessary dataset. Each video clip has an average duration of ten seconds portraying non-overlapping vehicle passages.

Duration	Clips Attained
17	56
29	64
16	44
22	25
34	79
46	88
53	96
28	34
32	68
11	76

Table 1: Video Clips to Form a Dataset

c. The attained video clips are manually annotated to detect motorcycles and an object detection boundary box is placed over each concerned object in the frame, as shown in figure 3. With a classifier identifying a motorcycle and helmet in different frames, our dataset collection is complete.

Figure 3: Annotated Video Clips

The dataset is categorized into two parts weighing 80% of the entire collection for training and the rest is utilized for testing.

Step 2:

a. LinearSVC model is chosen as our image classification depiction. It determines a hyperplane which categorizes whether the object in the frame is a motorcycle or not. Likewise, a separate LinearSVC model is developed using the helmet dataset to classify a helmet.

b. The inputted video data is collectively taken as frames. Each frame is a combination of several pixels. The image is identified through its pixels. These pixilated forms of numerical demonstrate a two-dimensional NumPy array. Through various feature extraction parameters such as spatial _feat and hog channel, our model is fed the data for classification predictions.

c. With LinearSVC's multi-class strategy, the formulated multi-class SVM determines the necessary co-efficient and weights (sample_weight parameter) of the model to increase the accuracy rate of its predictions. Predictions are represented in tables 2 and 3. The model training parameters are shown in figure 4.

Vehicle Classifier	Object	Prediction Value
	Motorcycle	0
	Non–Motorcycle	2

Table 2: Vehicle Classifier Prediction Labels

Helmet Classifier	Object	Prediction Value
	Helmet	0
	No Helmet	1

Table 3: Helmet Classifier Prediction Labels

```
LinearSVC(C=1.0, class_weight=None, dual=True, fit_intercept=True,
     intercept_scaling=1, loss='squared_hinge', max_iter=1000,
     multi_class='ovr', penalty='l2', random_state=None, tol=0.0001,
     verbose=0)
```

Figure 4: LinearSVC Model Parameters

d. Squared Hinge Loss is used to detecting the amount of deviation procured in the results from the original dataset values. Figure 5 demonstrates the accuracy score attained for the model trained.

precision	recall	f1-score	support
0.876	0.9764	0.9764	5882

Figure 5: Precision, Recall, F1-Score, Support

e. Evaluation is attained through the ratio of correct number of predictions made by the model to the total number of predictions made.

Step 3:

a. The trained pickled model is loaded through joblib.load() into our desktop interface application.

b. HOG Descriptors are used for foreground identification of the necessary object in particular frames. Resizing through pixels_per_cell and using transform_sqrt for power low compression, the object identification is done through feature extraction from the image data.

c. Keep track of the entire numerical data corresponding to the number of two-wheelers in the frame, motorcyclists with helmets, and motorcyclists without a helmet. Update the count to MySQL Database through mysql.connector.cconnect(), as shown in figure 6.

01:11:43	7	32
01:12:13	8	49
13:21:51	4	0
13:22:20	9	7
13:22:49	10	7
13:51:02	5	20
19:43:02	5	22
19:43:31	8	35
20:07:17	4	26
20:07:46	6	41
20:08:15	9	61
20:08:44	11	85

Figure 6: Count Stored in MySQL Database

d. Allowing access to only authorized users, the desktop interface has a login page through which you can gain access and live stream surveillance footage. Figures 7 and 8 visualize the login page and live to stream.

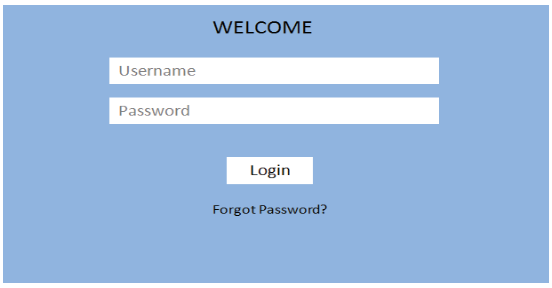

Figure 7: Login Page

Figure 8: Live Streaming Count

e. Through Pandas library, a tabular and graphical representation of the data attained is demonstrated as shown in figures 9 and 10.

Time Stamp	Helmet	No Helmet	Total
20-04-2020 01:11:43	7	32	42
20-04-2020 01:12:13	8	49	65
14-05-2020 13:21:51	4	0	6
14-05-2020 13:22:20	9	7	20
14-05-2020 13:22:49	10	7	21
04-06-2020 13:51:02	5	20	26
04-06-2020 19:43:02	5	22	29

Figure 9: Tabular Representation of Data Attained

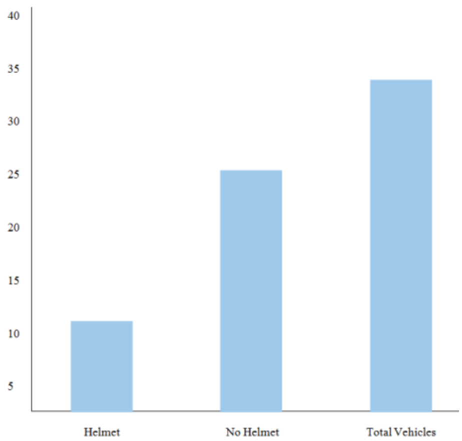

Figure 10: Graphical Representation of Data Attained

VI. CONCLUSION AND FUTURE WORK

Through this paper, a machine learning based approach is positioned to identify the helmet usage among motorcyclists. Video frames attained from surveillance footage, the object detection-based algorithm is trained to spot motorcycles and their helmet. Through various tools and methods corresponding to OpenCV and support vector classification, the desktop interface application is made possible to visualize the live streaming traffic surveillance footage. When trained, the LinearSVC algorithm acts in correspondence to the already existing road traffic surveillance and produces the corresponding data related to the usage of helmets owned by motorcyclists. The implemented algorithm has demonstrated high accuracy results and diverge slightly when the object is further away from the field. With 87.6% model accuracy rate, this research work has proposed a solution to enhance the driving safety measurements, which in turn deploys a time-efficient approach to handle the traffic regulations. An approach has to be proposed to identify the license number plate of each motorcyclist, who is disregarding the safety protocols in the future. The identified number plate will be updated and stored in the MySQL database to take further action on the riders. This method could eventually provide a time-and-effort efficient approach to tackle road safety abusers. Greater accuracy generating image classification and recognition algorithms such as CNN will prove to be more beneficial. The only subsidiary due to the usage of CNN is the requirement for larger data processing units and increases the corresponding time taken for training the model.

VII. REFERENCES

[1] Wonghabut, J Kumphong, T Satiennam, R Ung-arunyaweel, and W Leelapatra, “Automatic helmet-wearing detection for law enforcement using CCTV cameras”. IOP Conf. Series: Earth and Environmental Science 143 (2018) 012063

[2] e Silva, Romuere Rodrigues Veloso, Kelson Rômulo Teixeira Aires, and Rodrigo de Melo Souza Veras. "Helmet detection on motorcyclists using image descriptors and classifiers." *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*. IEEE, 2014

[3] N. Boonsirisumpun, W. Puarungroj and P. Wairotchanaphuttha, "Automatic Detector for Bikers with no Helmet using Deep Learning," *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, Thailand, 2018, pp. 1-4, doi: 10.1109/ICSEC.2018.8712778.

[4] Shine, L., C. V., J. Automated detection of helmet on motorcyclists from traffic surveillance videos: a comparative analysis using hand-crafted features and CNN. Multimed Tools Appl 79, 14179–14199 (2020).

[5] Kunal Dahiya, Dinesh Singh, C. Krishna Mohan, “Automatic Detection of Bike-riders without Helmet using Surveillance Videos in Real-time”. Conference Paper · July 2016 ,DOI: 10.1109/IJCNN.2016.7727586

[6] K. Dahiya, D. Singh, C.K. Mohan, “Automatic detection of bike-riders without helmet using surveillance videos in real-time”. International Joint Conference on Neural Networks (IJCNN), IEEE (2016), pp. 3046-3051

[7] M .C. Fong, J.R. Measelle, J.L. Dwyer, Y.K. Taylor, A. Mobasser, T.M. Strong, S. Werner, Ouansavanh, A. Mounmingkham, M. Kasuavang, et al. “Rates of motorcycle helmet use and reasons for non-use among adults and children” in Luang Prabang, Lao People's Democratic Republic BMC Public Health, 15 (1) (2015), p. 970

[8] J. Li et al., "Safety helmet wearing detection based on image processing and machine learning," 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), Doha, 2017, pp. 201-205, doi: 10.1109/ICACI.2017.7974509.

[9] C. Chiu, M. Ku and H. Chen, "Motorcycle Detection and Tracking System with Occlusion Segmentation," Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '07), Santorini, 2007, pp. 32-32, doi: 10.1109/WIAMIS.2007.60.

[10] J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," in IET Intelligent Transport Systems, vol. 6, no. 3, pp. 259-269, September 2012, doi: 10.1049/iet-its.2011.0138.

[11] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.

[12] C. Vishnu, D. Singh, C. K. Mohan and S. Babu, "Detection of motorcyclists without helmet in videos using convolutional neural network," 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, 2017, pp. 3036-3041, doi: 10.1109/IJCNN.2017.7966233.

[13] 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.

[14] Ramirez-Quintana, Juan Alberto, and Mario Ignacio Chacon-Murguia. "Self-adaptive SOM-CNN neural system for dynamic object detection in normal and complex scenarios." *Pattern Recognition* 48.4 (2015): 1137-1149.

[15] Y. Kulkarni, S. Bodkhe, A. Kamthe and A. Patil, "Automatic number plate recognition for motorcyclists riding without helmet," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, 2018, pp. 1-6, doi: 10.1109/ICCTCT.2018.8551001.