ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2021.0337

A BEHAVIORAL APPROACH TO DETECT SOMNOLENCE OF CAB DRIVERS USING CONVOLUTIONAL NEURAL NETWORK

V. Vinodhini, M. Abishek and K. Divya
Department of Computer Science and Engineering, Sona College of Technology, India

Abstract

The Road Traffic Accident Statistics concluded that fatal road accident 60% is caused by vehicle collision of taxi drivers. World Health Organization (WHO) is constantly initiating global road safety measures to minimize road accidents but, the cause for fatal injuries is primarily due to driving fatigue. Most people rely on cabs as the main transport. To provide obligate care of passengers, a computer vision-based technique is needed to detect the somnolence of drivers. Our proposed model CabSafety is based non-intrusive computer vision technique using Convolutional Neural Network (CNN). A tiny camera is fixed focusing the driver's face to monitor the behavioral changes like an eye blink, yawing, watery eye, mouth movement, and head position. The measures of the driver's eye are concentrated to identify sleepiness under stimulator or test conditions. The efficiency of the proposed model provides better results compared to the existing technique. The image from the camera is processed by OpenCV and Keras/Tensor flow. CNN classifier is used to detect eye status. The prediction from the CNN classifier produces an alarm to alert the driver.

Keywords:

Road Accident, Drowsy Driver, Eye Tracking, Convolutional Neural Network

1. INTRODUCTION

Driving a car is important to people in general because it offers an opportunity for personal control and autonomy. Nowadays, most people prefer the cab to the bus and the auto rickshaw, since long-distance travel by the cab would be more convenient for the passengers. According to the traffic department, about 60 percent of the road is occupied by taxis, the majority of taxi drivers' work throughout the day to make extra incentives, even if that means staying at the wheel at the cost of their sleep. Exhausted and drowsy taxi drivers are becoming a prime cause for accidents in the city. Cab drivers intend to earn extra money, accept rides at night too, after a hard day at work. If they're tired in the night ride, they pay no attention to the road that may cause an accident. Driver somnolence is the main cause of a road traffic accident. Driver sleepiness detection is very important in car safety technology to prevent road accidents. The U.S. National Highway Safety Agency (NHTSA) estimates that drowsy driving has resulted in nearly 100,000 road accidents and more than 1,500 deaths each year [1].

In our proposed system, behavioral interventions are used to identify sleeping drivers. Several face recognition systems were used to detect the face from the input image. Detection of the human face is simpler, but this is complicated in computer vision [2]. Face Detection Techniques are divided into functionality and images related techniques. Statistical, Neural Networks and Liner subspace methods have been used for image-based facial recognition approaches. Various eye region detection algorithms have been used to identify and remove the eye region from human face pictures [3]. Following the position of facial regions, normalization takes place in the pre-processing stage to eliminate the effects of illumination. The benefit of driver sleepiness is that it protects the driver without sleeping through the driver sleepiness monitoring system. The accuracy was improved by the use of facial landmarks identified by the camera and transferred to the Convolutional neural network (CNN).

2. LITERATURE SURVEY

Deng [4] proposed Dricare model to detect fatigue status of driver. The parameters considered for fatigue are yawning, duration of eye closure and blinking. 68 Facial key points were considered for fatigue detection model. A camera is mounted to capture the video of the driver. Multiple Convolution Neural Network (MCNN) with Kernelized Correlation Filters (KCF) algorithm uses manual method to consider the drivers face. To improve the accuracy he designed a face tracking algorithm. A technique called electrooculogram (EOG) signal and blinking feature is used to detect drowsiness. The Dricare model is based on video processing that is shot with a camera and processed in the cloud. Facial key points are recognized by the Cloud algorithm. The cloud server senses the driver status. The results are sent as a warning to the deriver's cellphone.

Stern [5] proposed various causes for drowsiness of drivers using multivariate approach for real time ambulatory drivers which leads to accidents. The technologies used to record the movements of drivers are electronic data recorders, actigraphy, and mobile phone. It also considers the features, sleeping habits, medication, long-term health effects and driving hours of the car. The motion of the car, the number of times the brakes are applied and the state of the car is often taken into account. The driving length of the drivers is known to be the key cause of the road accident. The features of the environment like weather, rainfall, speed limits, vehicle traffic, design and safety measures are vital to the possibility of a crash. In measuring the relationship among fatigue drivers and crashes method is proposed by real-time assessment.

Schwartz [6] proposed a driver monitoring system to identify the sleepiness of the driver. The author associated their work with the existing work with sensors associated with vehicle. The experiment included twenty drivers on duty in high fidelity simulators for four hours in a Controlled Area Network (CAN). These drivers controlled the vehicle in slow and faster rates. Drowsiness is categorized into two types like micro sleep and fatigue state. The blink, eye movements are predicted in micro sleep category. The ground truth 0 to 4 range is not considered to be drowsy rather scores greater than 4 detects sleepy state. The results with different measures are obtained. Driver monitoring system combined both driver signs and vehicular signs which gave better prediction than the existing models.

Ramzan [7] proposed relevant factors that signs for accidents. The features include facial movements, biological characteristics of drivers. He analyzed existing models and classified the techniques applicable to resolve the crashes. The data gathering is done from various case studies, previous research on drowsiness and its results are highlighted under various criteria. The Yawning extraction with eye blink monitoring, eye closure monitoring are categorized under behavioral method. Lane detection, steering angle is under vehicular parameter method. The techniques based on electroencephalogram, wavelet analysis with heart rate, Pulse detection with sensor are under psychological based technique. The classification of error rate and accuracy for various models are given.

3. CAB SAFETY MODEL

Traditional methods are not capable to detect the facial expression accurately. This paper proposes a Cab Safety Device which is represented in Fig.1. It contains the CNN algorithm and illumination algorithm that detects the drowsiness of the driver that alert with an alarm. The real time data are stored in cloud server.

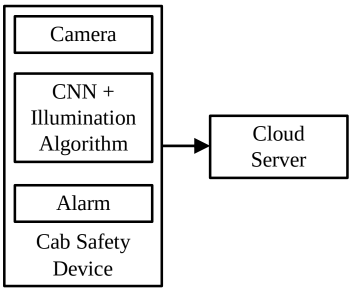


Fig.1. Cab Safety Device

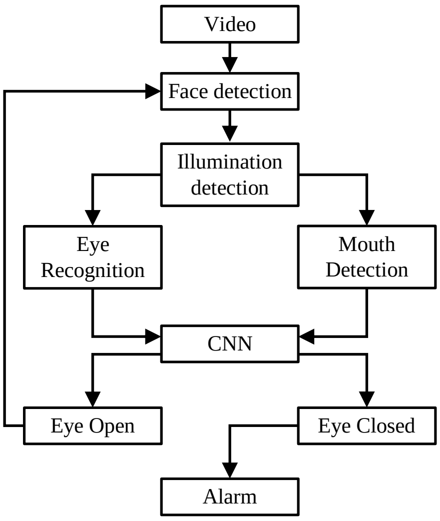


Fig.2. Flow Diagram of Proposed Model

The principle workflow of the Cab Safety device is shown in Fig.2. The cab device has an inbuilt dash cam in it and an alarm to indicate the driver on fatigue condition. The video captured is processed with 25 frames per second. The CNN detects face and Adaptive Histogram Equalization (AHE) algorithm is used to enhance the quality of the image. It improves the contrast and adjusts the brightness of the input image [8]. The Fig.4 and Fig.5 represents the original input image and the enhanced output image. This algorithm guarantees accuracy in such a way that the prediction is enhanced by conventional channels.

The parameter to detect the drowsiness is given in Fig.3. The evaluation is primarily based on eye and the mouth. The eye and mouth is predicted by CNN algorithm. Convolution Neural Network (CNN) comes under Neural Network it has 1 or more layers and it's primarily used for image processing, classification and identification. Identification of a text and categorizing it, Digital signature and Google Lens can be considered as an example. It is also used in NLP (Natural Language Processing) for speech recognition and speech synthesizer. Each layer in CNN is consists of a chain of filters called as convolutional kernels. It is a matrix of numbers that are in a subset of input pixels. Every pixels are multiplied with the value of the kernel and the obtained value gets summarized to a single value as output as feature map.

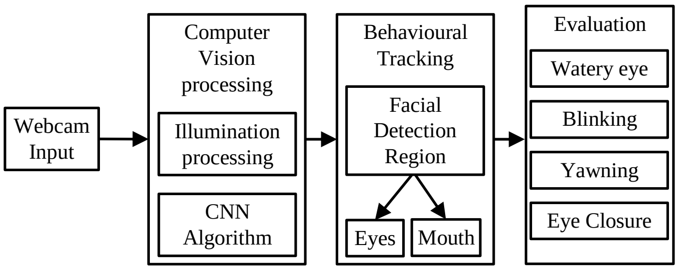


Fig.3. Block Diagram of proposed model

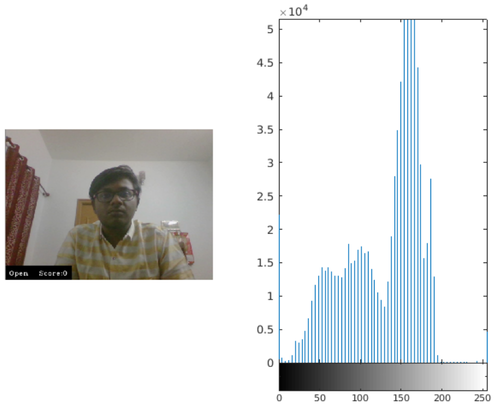


Fig.4. Original Image

To enhance the image quality, Adaptive Histogram Equalization algorithm (AHE) is used [9]. The Fig.4 depicts the histogram of the captured image. After applying the AHE (detects the luminous to be enhanced), the Fig.5 shows the enhanced image. The AHE technique includes Num Tiles, Clip limit, Normalization and probability, where Num Tiles is [2,2] the region to be covered. Clip Limit is set to 30. The grayscale image is again converted to RGB and fed to CNN.

Convolution of an image with a variety of filters can perform many tasks such as image illumination, edge detection and dynamic color enhancements [10]. At few circumstances the filter won't get fit perfectly with the image, at these situations we either need to surround the image with 0 or to delete the parts of the image where it doesn't get fit. These are known as padding. A stride defines number of pixels that moves filters to number of pixels at a time. For example if it is 3 then we move in filter 3 pixels at a time.

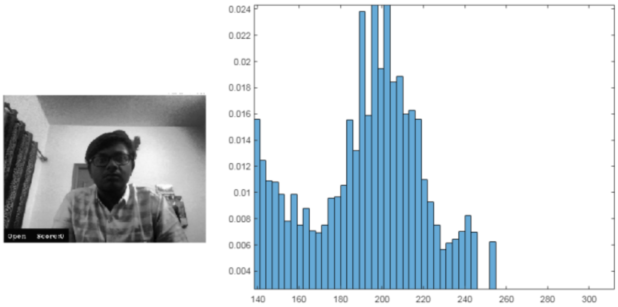


Fig.5. Output of Adaptive Histogram Equalization

Normalization is technique to change the numerical columns of a dataset to a common scale, without varying or false difference in the ranges of the value in order to find optimal parameters [11]. It can be obtained by subtracting the mean and dividing it with its standard deviation. The data scaling of images ranges 0 to 1. It can be just by dividing the pixel values with 255.

Nonlinearity is a situation where there is no direct relationship between the dependent and independent variables (No straight line) [12]. Rectified Linear Unit (ReLU) is used to bring non-linearity to our neural network to avoid negative values by converting these values to 0. The output of the ReLU is

$$f(x)=\max(0,x).$$

Pooling layer is used to summarize the features of feature map to patches. There are 3 types of pooling method.

- *Max Pooling*: It takes the largest element from the feature map
- *Average Pooling*: It summarizes the average presence of an element from the feature map.
- *Sum Pooling*: It sums up all the elements from the feature map.

The algorithm for somnolence detection of cab driver (given below) is calculated based on different parameters [13]. The parameters like eye closure, closed eye with yawning, the eye blink ratio, yawning ration these parameters are validated with a threshold value. The total score predicts the driver fatigue ratio.

Algorithm 1: Somnolence Detection of Cab Driver using Cab Safety

Input: Frames captured from the cab camera

Output: Predicting whether the driver is fatigue or awake displayed via score

Loads the frames (image) of video

To detect the Eye Closure ratio

If c>36% then

  W_c= 1

End if

// Scans the eyes and mouth to detect their status

//C is the ratio of eye closure in a minute

To detect the Eye Closed Time and Yawning

If t>5s and is not yawning then

  W_t = 1

End if

//t represents how long the eyes were closed (duration) in a minute

To detect the Blink Ratio

If b>30 or b<3 then

  W_b=1

End if’

//b represents the number of blinking of the driver in a minute

To detect Yawning Ratio

//y represents the number of yawning in a minute

If y>5s then

  W_y= 1

End if

Total Weight Calculation

//W represents weights of each parameter

//S represents the score

Calculate S the total value of these weight.

$$S=W_c+W_t+W_b+W_y$$

If S≥3 then

  The Driver is predicted to be fatigue

Else

  The driver is not asleep

End if

The proposed model reaches an accuracy of 96% (approximately) when the Euclidean distance of eyes are within 23px.

Table.1. Comparison with other Models

Algorithm	Frames Per Second (FPS)	Accuracy (%)
MTCNN	3	93.2
KCF	188	91
DSST	6	85
Struck	8	76
KCF+CNN	26	93
MC-KCF	25	95
CNN + Image	24	96
Cab Safety	25	97

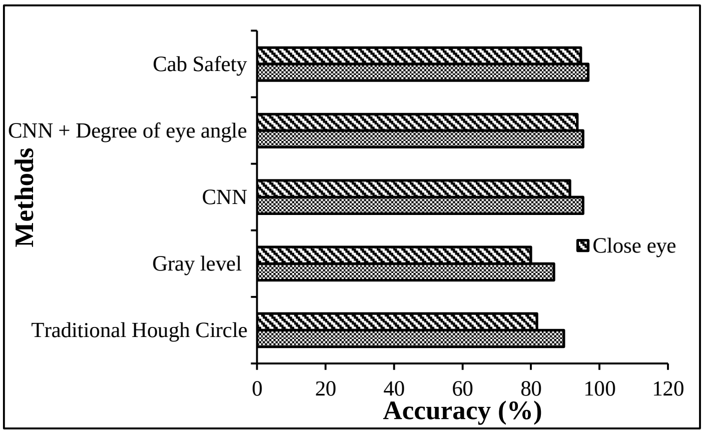


Fig.6. Comparison of Eye Motion

From Table.1 it is clear that our algorithm (Cab Safety) is much efficient in terms of predicting and has a decent frame rate of 24 Frames Per Second (FPS). Therefore our algorithm offers better than many of the other algorithms in terms of accuracy.

The Fig.6 depicts the comparison of eye motion of different methods based on their eye ball speed. The eye recognition is the highest priority in drowsy detection among all parameter. The success rate of identifying closed eye is nearly 94.6%.

4. RESULTS AND DISCUSSION

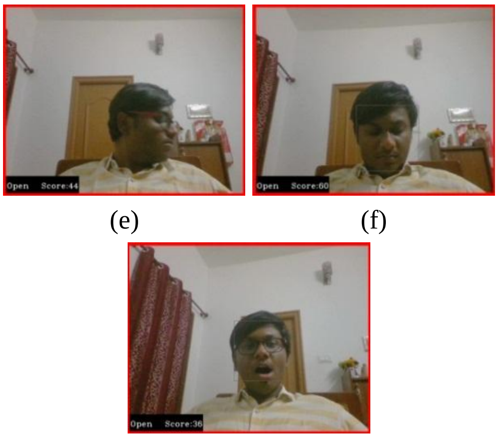
The experiment is carried out under different environmental conditions from Table.2, it is understood that the proposed system has the best accuracy even when the driver wears glasses. It produces good results for drivers in a good lightning environment. The illumination algorithm pays a vital role for the drivers under bad lightning conditions [14].

The adaptive histogram equalizer algorithm thereby increases the contrast of the image. However there are quality loss and accuracy falls down if he/she wears glasses since we couldn’t obtain the complete data of eye and also the frame drastically drops to 16FPS since this algorithm needs to be implemented each every frame of the image [15].

Table.2. Driving in different Environmental Status

Driving Environmental Status	Detection Rate (%)	Accuracy (%)
Driver with Spectacles	92	91
Driver without Spectacles	96.7	95
Driver with Spectacles and good lighting	92.4	90
Driver without Spectacles and good lighting	96.7	96
Driver with Spectacles and Dim light	90.5	89.2
Driver without Spectacles and Dim light	93	92
Yawning in good lighting	92	91.6
Yawning in bad lighting	89	87.8

(a) (b) (c) (d)



(e) (f) (g)

Fig.7. Experiments under Various Parameter (a) Original Image (b) Eye closed with head down (c) Eye Closed (d) Yawning (e) Turned towards left (f) Without glass (g) Opened mouth

5. CONCLUSION

The cab safety model is a real-time model, evaluates the somnolence of cab drivers based on behavioral approach. The parameter of behavioral approach includes eye movements, eye closure, yawning, and eye blink. A new algorithm for cab safety is proposed based on the parameters along with Adaptive Histogram Equalizer (AHE) to enhance the contrast of the image and CNN model to predict the fatigue condition. The experiment is carried out for all possible conditions at various environments. The detection rate and accuracy is compared with the existing methods proved to be stable. The proposed method provides better performance and efficiency even if the driver wears glasses.

REFERENCES

[1] R. PrietoCuriel, H. Gonzalez Ramirez and S.R. Bishop, “A Novel Rare Event Approach to Measure the Randomness and Concentration of Road accidents”, *PloS One*, Vol. 13, No. 8, pp. 1-12, 2018.

[2] A. Rossler, D. Cozzolino, L. Verdoliva and J. Thies, “Faceforensics: A Large-Scale Video Dataset for Forgery Detection in Human Faces”, *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 1-12, 2018.

[3] A.S. Zandi, A. Quddus and L. Prest, “Non-Intrusive Detection of Drowsy Driving based on Eye Tracking Data”, *Transportation Research Record*, Vol. 2673, No. 6, pp. 247-257, 2019.

[4] W. Deng and R. Wu, “Real-Time Driver-Drowsiness Detection System using Facial Features”, *IEEE Access*, Vol. 7, pp. 118727-118738, 2019.

[5] H.S. Stern, D. Blower and M.L. Cohen, “Data and Methods for Studying Commercial Motor Vehicle Driver Fatigue, Highway Safety and Long-Term Driver Health”, *Accident Analysis and Prevention*, Vol. 126, pp. 37-42, 2019.

[6] C. Schwarz, J. Gaspar and T. Miller, “The Detection of Drowsiness using a Driver Monitoring System”, *Traffic Injury Prevention*, Vol. 20, No. 2, pp. 157-161, 2019.

[7] M. Ramzan, H.U. Khan and S.M. Awan, "A Survey on State-of-the-Art Drowsiness Detection Techniques", *IEEE Access*, Vol. 7, pp. 61904-61919, 2019.

[8] G.D. Finlayson, "Colour and Illumination in Computer Vision", *Interface Focus*, Vol. 8, No. 4, pp. 1-12, 2018.

[9] J. Ma, X. Fan and S. Yang, "Contrast Limited Adaptive Histogram Equalization-Based Fusion in YIQ and HSI Color Spaces for Underwater Image Enhancement", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 32, No. 7, pp. 1-14, 2018.

[10] Y. Ji, S. Wang and Y. Zhao, "Eye and Mouth State Detection Algorithm based on Contour Feature Extraction", *Journal of Electronic Imaging*, Vol. 27, No. 5, pp. 1-13, 2018.

[11] J. Ou and Y. Li, "Vector-Kernel Convolutional Neural Networks", *Neurocomputing*, Vol. 330, pp. 253-258, 2019.

[12] A. Khan, A. Sohail and U. Zahoora, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks", *Artificial Intelligence Review*, Vol. 23, No. 1, pp. 1-16, 2020.

[13] C. J. Bourdin and J.L. Vercher, "Detection and Prediction of Driver Drowsiness using Artificial Neural Network Models", *Accident Analysis and Prevention*, Vol. 126, No. 1, pp. 95-104, 2019.

[14] S. Sahu, A.K. Singh and S.P. Ghrera, "An Approach for De-Noising and Contrast Enhancement of Retinal Fundus Image using CLAHE", *Optics and Laser Technology*, Vol. 110, pp. 87-98, 2019.

[15] V.B. Hemadri, P. Gundgurti and K. Deepika, "A Novel on Biometric Parameter's Fusion on Drowsiness Detection using Machine Learning", *Proceedings of International Conference on Computer Networks and Communication Technologies*, pp. 1-12, 2019.