International Journal of Advanced Development in Science and Technology
Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

A MULTITASK CNN FRAMEWORK FOR DRIVER FATIGUE DETECTION

[1] Nandha Krishnan M R, [2] RamKumar S, [3] Umamaheshwari
[1,2] Student, [3] Supervisor,
Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar,
ramkumarsankar99@gmail.com.

ABSTRACT

Driving fatigue is one of the most important factors in traffic accidents. In this paper, we proposed an improved strategy and practical system to detect driving fatigue based on machine vision and Adaboost algorithm. Changes and progresses in information technologies have played an important role in the development of intelligent vehicle systems in recent years. The System detects the drivers' fatigue status, such as yawning, blinking, and duration of eye closure, using video images, without equipping their bodies with devices. Owing to the shortcomings of previous algorithms, we introduce a new face-tracking algorithm to improve the tracking accuracy. In this study, the fatigue degree of the drive is divided into 3 classes. The proposed model achieved 98.81%fatigue detection on YawDD and NthuDD dataset. The success of the model is presented comparatively.

Keywords:Convolutional neural network, driver fatigue detection, PERCLOS, FOM.

I. INTRODUCTION

According to World Health Organization (WHO) statistics, traffic accidents cause millions of people to lose their lives every year. Statistics assert that most of the fatal accidents are due to driver fatigue and carelessness. The American Automobile Association reported that 7% of all accidents and 21% of fatal traffic accidents were caused by tired drivers. The US National Highway Traffic Safety Administration (NHTSA) states that only 2.2% to 2.6% of total annual fatal accidents in the USA between 2005 and 2009 stemmed from driver fatigue. Accidents that caused only material damage are not included in these results. According to reports in 2009, about 30,000 injury accidents (2.0% of all injuries in 2009) stemmed from driver fatigue. In 2017, Foundation for Traffic Safety's study found that in a normal week, 42.4% of drivers were driving without at least one or more days of sleep, less than six hours of sleep, and were serious for the majority of drivers (87.9) and they perceive what they see as unacceptable behaviour (95.2%). However, about 3 out of 10 drivers (30.8%) admit that they drove cars even though they were too tired to keep their eyes open in previous months.

The process of falling asleep in the vehicle is a gradual event. Due to monotonous driving conditions and other environmental factors, the driver can change from normal to drowsiness. Therefore, the first critical issue to be identified in the fatigue detection system is how to detect drowsiness accurately and early. In this article, it is recommended to monitor driver fatigue in real time using a behavioural model on drivers. Behavioural fatigue detection methods can be applied without distracting the driver and can capture this gradual transition.

There are currently four methods of driver fatigue detection:

1. Based on physiological indicators, Rohit et al. analysed the characteristics of electroencephalogram (EEG) using linear discriminant analysis and a support vector machine to detect driver fatigue in real-time. However, most on-board physiological sensors are expensive and must be attached to human skin, which can cause driver discomfort and affect driver behaviour.

2. Based on the driving state of the vehicle, Ramesh et al. used a sensor to detect the movement state of the steering wheel in real time to determine the degree of driver fatigue. However, the primary disadvantage of this method is that detection is highly dependent on the individual driving characteristics and the road environment. As a result, there is a high degree of randomness and contingency between the driving state of the vehicle and driver fatigue, which reduces the detection accuracy.

3. Based on machine vision, Grace measured pupil size and position using infrared light of different wavelengths. Yan et al. used machine vision to extract the geometric shape of the mouth shape. An advantage of this method is that the facial features are non invasive visual information that is unaffected by other external factors, i.e., driving state of the vehicle, individual driving characteristics and road environment.

4. Based on information fusion, Wang Fei et al. combined physiological indicators and driving state of the vehicle to detect the driver fatigue state of the driver by collecting the EEG signal of the subject and the corresponding steering wheel manipulation data. However, the robustness of the test is affected by the individual's manipulation habits and the driving environment.

II. RELATEDWORK

A. Visual Object Tracking

Visual object tracking is a crucial problem in computer vision. It has a wide range of

International Journal of Advanced Development in Science and Technology
IJADST Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

applications in fields such as human-computer interaction, behaviour recognition, robotics, and surveillance. Visual object tracking estimates the target position in each frame of the image sequence, given the initial state of the target in the previous frame. Lucas and Kana de proposed that the tracking of the moving target can be realized using the pixel relationship between adjacent frames of the video sequence and displacement changes of the pixels. However, this algorithm can only detect the medium-sized target that shifts between two frames. With the recent advances of the correlation filter in computer vision, Bolme proposed the Minimum Output Sum of Squared Error (MOSSE) filter, which can produce stable correlation filters to track the object. Although the MOSSE's computational efficiency is high, its algorithm precision is low, and it can only process the gray information of a single channel.

Based on the correlation filter, Li and Zhu utilized HoG, color-naming features and the scale adaptive scheme to boost object tracking. Danelljanet al. used HOG and the discriminative correlation filter to track the object. SAMF and DSST solve the problem of deformation or change in scale when the tracking target is rotating. Further, they solve the problem of the tracker's inability to track object adaptively and the low operation speed. With the development of the deep-learning

algorithm, some scholars combine deep learning and the correlation filter to track the mobile target. Although these algorithms have better precision than the track algorithms based on the correlation filter, their training is time-consuming. Hence, these algorithms cannot track the object in real-time in a real environment. In this study, we propose a MC-KCF algorithm based on the correlation filter and deep learning. This algorithm uses CNN and MTCNN to offset the KCF's limitation and uses the KCF to track objects. Thus, the algorithm can track the driver's face in real-time using our system.

B. Facial Landmarks Recognition

The purpose of facial key-points recognition is that getting the crucial information about locations of eyebrows, eyes, lips and nose in the face. With the development of deep learning, itis the first time for Sun et al. to introduce DCNN based on CNN to detect human facial key points. This algorithm only recognizes 5 facial key points, albeit its speed is very fast. To get a higher precision for facial key points recognition, Zhou et al. employed FACE++ which optimizes DCNN and it can recognise 68 facial key points, but this algorithm includes too much of a model and the operation of this algorithm is very complicated. Wu et al. proposed Tweaked Convolutional Neural Networks (TCNN) which is based on Gaussian Mixture Model (GMM) to improve different layers of CNN. However, the robustness of TCNN depends on data excessively. Kowalski et al. introduced Deep Alignment Network (DAN) to recognize the facial key points, which has better performance than other algorithms. Unfortunately, DAN needs vast models and calculation based on complicated functions. So in order to meet the requirement about real time performance, DriCare uses Dlib to recognise facial key points.

C. Driver Drowsiness Detection

Driver drowsiness detection can be divided into two types: contact approaches and non-contact approaches. In contact approaches, drivers wear or touch some devices to get physiological parameters for detecting the level of their fatigue. Warwick et al. implemented the Bio Harness 3 on the driver's body to collect the data and measure the drowsiness. Li et al. used a smartwatch to detect driver drowsiness based on electrodense philographic(EEG)signal. Jung et al. reformed the steering wheel and set an embedded sensor to monitor the electrocardiogram (ECG) signal of the driver. However, due to the price of contact approaches and installation, there are some limitations which cannot be implemented ubiquitously. The other method employs a tag-free approaches to detect the driver drowsiness, where the measured object does not need to contact the driver. For example, Omidyegan eh et al. used the driver's facial appearance captured by the camera to detect the driver drowsiness, but this method is not real-time. Zhang and Hua used fatigue facial expression reorganization based on Local Binary Pattern (LBP) features and Support Vector Machines (SVM) to estimate the driver fatigue, but the complexity of this algorithm is bigger than our algorithm. Moreover, Picot et al. proposed a method that uses electrooculogram (EOG)signal and blinking feature for drowsiness detection. Akrout and Mahdi and Oyini Mbounaet al. used a fusion system for drowsiness detection based on eye state and head position. Different from these methods, we employ simple formulae and evaluations, which make the results easier to measure.

III.PROBLEM DEFINITION

In this study, driver fatigue detection is performed with Multi-task using raw data obtained from different data sets in the literature. In recent driver fatigue detection systems, most studies have focused on using limited visual cues. However, human fatigue is a complex mechanism and depends on the dynamic cohesion of various cues, which means that outcomes and situations can be improved. Fatigue is a condition that requires continuity, and instant decisions cannot be made while driving. Fatigue at a predetermined time point is considered a factor for fatigue at the present time point, and time varies according to the behaviour of individuals. For example, yawning is a major symptom of fatigue, but it does not always occur before the driver sleeps. This should be considered as a preliminary step during fatigue detection and should be evaluated in consideration of the eye conditions of the driver. Otherwise, if something is not detected for a while, the probability of fatigue may be calculated incorrectly.

D. PRE-TRAINED STAGE

Dlib, a platform-independent programming library created in the C ++ programming language, contains this face pointer data set created by Sagonas et al. The Dlib Library's pre-trained facial landmark detector is used to predict the location of 68x-y coordinates that map facial landmarks in the facial zone. Detecting facial landmarks is a critical subject in terms of facial zone shapes estimation. In this study, the dlib library is used to detect and track the faces of the drivers in real time videos. Therefore, important facial structures are detected on the face zone using shape estimation methods. In order to be used in the Multi-Task CNN model, YawDD and Nthu-DDD video datasets are first resized from 640 ×480 resolution to 320 × 240. Then, face, mouth and eye regions were determined by using Dlib algorithm on these resized images. It is determined whether the mouth is open / closed and the eye is open / closed on these identified areas and the opening is labelled according to the closed state. The opening states are labelled as ‘‘1’’, and the closed states as ‘‘0’’. When training data is being prepared, the videos include blinking for the eye or opening before closing the eye, or looking at other angles. If 80% of the eye opening is open when labelling the data, the label value is considered ‘open’ and in the remaining cases itis considered ‘closed’. Likewise, during speech or opening the mouth at random situations are different from yawning. If 80% of the mouth opening is open, label value is considered ‘open’ and in other cases ‘closed’ in order to prevent errors during detection and to label these situations correctly. In each training, 20% of the training data is chosen as random verification data. The selection of verification data is consistent with the dataset distribution.

E. FATIGUE DETECTION

Fatigue parameters based on face markers:1)PERCENTAGE EYE CLOSURE (PERCLOS) PERCLOS is the percentage of

squares that appear to be closed to the total squares of the human eyes
in a given time frame. In other words, PERCLOS can be defined as a fatigue analysis method that shows the ratio of closed eyes depending on the number of open and closed eyes. This value can be calculated as in (1);

fPERCLOS =nclose/NCloseandOpen×100% (1)

NClose and Open represents the total number of open and nclose represents closed eye frames at a given time and represents the number of closed-eye frames at a close, specific time. In the literature it has been determined that a driver blink approximately 10 times per minute under normal conditions. Higher low PERCLOS is a method used in the literature to detect eye fatigue. 2) FREQUENCY OFMOUTH(FOM) FOM is the ratio shown as a percentage of squares that open to total squares in a given time frame. Calculation of FOM (2) is similar to PERCLOS calculation.
fFOM = n open/NCloseandOpen×100% (2)

n open is the number of open mouth frames in a period, NClose and Open represents the total number of frames in a period.
IV.PERFORMANCE METRIC

Two different dataset (YawDD and NthuDDD) are used in the study for training and testing. With the dlib algorithm, the location of 68 points in the face region is captured on determination of the frequency range. Therefore, PERCLOS and FOM are calculated within the fixed the video to determine the driver eye and mouth regions. Thus, mouth and eye status information is labelled. Labelled data is classified by multi-task CNN. In order to calculate fatigue parameters accurately, a certain frequency range is required (total number of frames: N). The frequency range to be used here is kept constant within the study and the number of frames is fixed. Each new frame is removed and the last frame is removed. Thus, the frequency range remains constant. In the proposed study, during the training process, the data sets are divided into training set, validation set and test set. The distribution of the data set of the YawDD and Nthu-DDD architectures used in the multi-taskarchitectureisshowninTable1.
In the study, since both the eye and the mouth were simultaneously classified on the images during the training phase, the sunglass driver data contained in the YawDD and Nthu-DDD dataset was not used. Stochastic Gradient Decent (SGD)optimization algorithms are used in the training of the proposed model. SGD parameters batch size = 64, 0.6 momentum and 0.0005 weight distortion is used. The training speed is startedat0.01foralltrainablelayers.

When drivers fatigue, they will have series of behavioural reactions such as eyes closed or yawning. There- fore, it can be accepted, drier fatigue can be calibrated by calculating the PERCLOS and FOM parameter.

TABLE1.Distrubution of dataset

Dataset	YawDD dataset	Nthu-DDD dataset
Training	41600	23943
Validation	10800	4964
Test	10910	5000
TOTAL	**62910**	**33907**

FIGURE1. Proposed model of driver fatigue detection system.

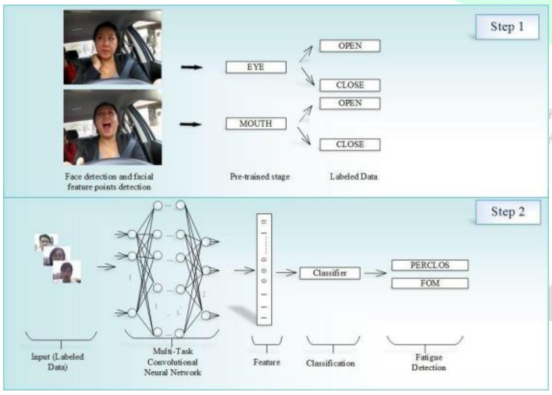
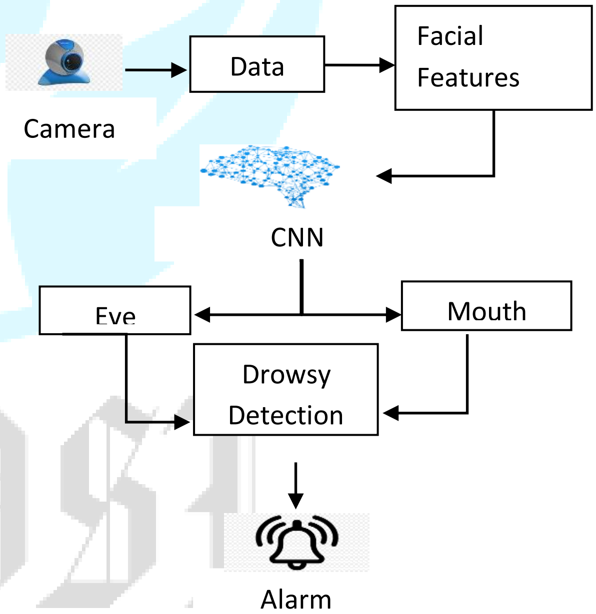


TABLE2.Multitask accuracy compared with other approaches.

	Year	Dataset	%
Zhang et al [9]	2015	YawDD	92%
Gu et al.[21]	2018	EMD,	96.89%
Ji et al. [21]	2019	YawDD	98.42%
Proposed apporach	2021	YawzW, NthuD	**98.81%**

V. ARCHITECTUREDIAGRAM



VI.CONCLUSION

In this article, Multi-task CNN models is used to detect driver fatigue in real time. The Dlib algorithm is used to accurately identify the driver's eye and mouth information. Then, the system is trained with Multi-task CNN models for the determination of fatigue parameters. The frequency range to be used here is kept constant within the study and the number of frames is fixed. Finally, depending on fatigue parameters, fatigue is evaluated as ''very tired, less tired and not tired''. These situations are also dynamically tested and coded at certain periods that maintain their continuity. The accuracy performance of the system tested in real time is very robust. The proposed system can model the interactive relationship between eye, mouth and sub-states. Fatigue at a predetermined time point is considered a factor for fatigue at the present time point, and time varies according to the behaviour of individuals. The system runs successfully. One of the most powerful features of the study is that it is a faster and more powerful system with a single model without creating a separate CNN model with two different architectures. In our future work, the head condition, which is as important as the eye and mouth condition, will be added to the system and the system will be integrated in to an embedded system.

REFERENCES

[1] B. C. Teff (2016). Acute Sleep Deprivation and Risk of Motor Vehicle Crash Involvement. Accessed: Jun.12,2020.

[2] T.S. Facts. (2011). National Highway Traffic Safety Administration. Accessed: Sep.25,2019. [Online]. Available: http://www.americanindian.net/sleepstats.html

[3] T. Johnson. (2018).2017 Traffic Safety Culture Index. Accessed: Sep.25,2019. [Online].Available:https://aaafoundation.org/2017-traffic-safety-culture-index/

[4] Q. Wang, J. Yang, M. Ren, and Y. Zheng, ''Driver fatigue detection: A survey,'' in Proc. 6th World Congr. Intell. Control Autom. Jun.2006, pp.8587–8591.

[5] Y.Liu, H.Zhang, and J.Liu, ''Driver fatigue monitoring method based on eyes state classification,'' in Proc.IEEE Chin. Control Decision Conf.,Jul.2008,pp.2257–2260

[6] R. Coetzer and G. Hancke, ''Eye detection for a real-time vehicle driver fatigue monitoring system,'' in Proc.IEEE Intell.Vehicles Symp. (IV),Jun.2011,pp.66–71.

[7] B. K. Savaş and Y. Becerikli, ''Development of driver fatigue detection system by using videoimages,'' Akıllı Sistem

lerde Yenilikler Ve Uygula- maları, Konferansı-Bildiri Özetleri Kitabı, Alanya, Turkey,2017.

[8] X.-Q. Luo, R. Hu, and T.-E. Fan, ''The driver fatigue monitoring system based on face recognition technology,'' in Proc. 4th Int. Conf. Intell. Control Inf. Process. (ICICIP), Jun. 2013, Jun.2013.

[9] W. Zhang, Y. L. Murphey, T. Wang, and Q. Xu, ''Driver yawning detection based on deep convolutional neural learning and robust nose tracking,'' in Proc. Int. Joint Conf. Neural Network.(IJCNN),Jul.2015,pp.1–8.

[10] V. Le, J. Brandt, Z. L in, L. Bourdev and T.S. Huang, ''Interactive facial feature localization,'' in Proc. Eur. Conf. Comput. Vis. Berlin, Germany: Springer, Oct. 2012, pp. 679–692.