JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

A Visual Fusion Based Deep Learning approach for Real Time Driver Drowsiness Detection

I. Nikitha Satheesh[1] & Dr. John Basha[2]

[1]Department of Computer Science and Engineering,
Jain deemed-to-be University,
Bangalore, Karnataka, India.
Email : nikithasatheesh5319@gmail.com

[2]Professor,
Department of Computer Science and Engineering,
Jain deemed-to-be University,
Bangalore, Karnataka, India.
Email : *jjbasha@gmail.com*

ABSTRACT

Growing concerns due to increased road accidents due to driver drowsiness have paved the way for the development of real time driver sleepiness detection systems.The real world implementa -tion of these systems can play a major role in reducing the accidental rate due to driver sleepin -ess. This paper suggests a fusion based approach for identifying the driver's sleepiness in real-time.The model uses driver's behavioral features like eye state and yawning as a metric for sleepiness identification. Mouth Aspect ratio is used for yawn detection and Convolution Neural Network for eye state classification. Dlib's HOG based detection is used for facial feature detection in this paper. Our proposed model shows a high accuracy rate and can detect if the driver is drowsy or not. It performs well in various lighting conditions and also if the driver wears spectacles. Training and testing accuracy of the proposed CNN model is 98.75% and 97.65% respectively. Moreover, it is found that our proposed methodology performs well in real time.

Keywords: Drowsiness detection, CNN, Eye State, Yawning, Fusion based, MAR

1. Introduction

Everyone is familiar with the term "automobiles", its development has changed the transportation pattern of the world. It made life much easier to conduct daily activities.Automobiles are widely available in a wide variety ranging from fossil fuels to electric cars. Automobile use has costs and benefits. Some benefits of automobiles include flexibility and availability of transportation as and when needed, income generation and so on. Some costs of automobile use include air pollution, depreciation, vehicle disposal issues and traffic collisions. Studies reveal that deaths due to automobile collisions are very high around the world. For example, even though there were restrictions on vehicle movement due to covid induced lockdowns in India but still around 1.51 lakh in 2019 and 1.32 lakh in 2020 lives were lost due road accidents. Several research shows that 40% of road crashes are due to driver drowsiness.The correlation between road crashes and driver sleepiness can be found in various studies. Researchers have found that sleepy drivers are 4 -6 times more likely to be involved in a collision or near- collisions than attentive drivers and can be found that fatigue drivers showed more fast corrective steering wheel movements, larger trajectory deviations and less adherence to the speed limit.

Hence by reducing or eliminating collisions due to driver sleepiness can have a huge impact in the count of road accidents. Early detection of drowsiness is the weapon to avoid drowsy driver collisions. Various methods have been implemented to reduce road collisions ranging from self assessment to the use of computer vision techniques in drowsiness detection but still it does not completely eliminate driver drowsiness related collisions , so there is a need to build complementary warning models that can overcome the limitations of existing approaches.

The aim of our paper is to introduce a model that can detect and alert a drowsy driver. Our model will be using deep learning and computer vision techniques for early detection of driver sleepiness. As everyone knows eyes and mouth are the key indicators of drowsiness, we will be using these two indicators for driver sleepiness detection. The proposed approach is mainly divided into various steps like detection of face, eyes and mouth detection and extraction, fusion state and so on . The extracted eyes state is determined by CNN. If the driver is detected drowsy then an alert is sent to the motorist. This research paper is organized as : The 2nd section will take through a literature survey. The methodology is discussed in sections 3 and experiential results in section 4. Finally, section 5 contains a conclusion.

2. Literature Review

There are several methods to detect driver sleepiness and few of these methods will be discussed in this section. Parameters such as eye state, yawning, lane deviation, steering wheel , head bending , EEG and ECG can be used for detection of driver sleepiness. Computer vision and deep learning techniques mostly use eye state, yawning, lane deviation and so on for drowsiness detection. Camera is the key element of these techniques which is used for real time monitoring and image capturing.

Tanvir Ahammed [1] in his work considers sleepiness detection as detecting an object and uses MobileNet CNN with Single Shot Multibox Detection for sleepiness detection tasks.It considers yawn and eye state for driver fatigue detection. Jabbar [2] introduced a deep learning based android app for driver sleepiness detection. This method uses Dlib for facial coordinate extractions and these points are passed to the multi-layer perceptron classifier for classifying as drowsy or not drowsy.Jonathan ‘s [3]model extracts face region and passes to a custom designed shallow CNN (SS-CNN) to detect driver state using “eye closed” or “eye open”. So as to distinguish normal eye closeness due to eye blinking and drowsiness, this model also analyzes consecutive SS-CNN results. In the model proposed by R. Ayachi [4], EfficientDet-B0 is ude for driver sleepiness detection . Behavioral features such as eyes and mouth coordinates are detected and categorized as “open” or “close”. This model used the NTHU-DDD data set and achieved an average accuracy of 96.05% . Even though this model is fastest in object detection but requires a huge number of trainable parameters.

Jun-Juh Yan and others [5] in their research uses grayscale images to calculate the approximate position of the driver’s face and uses template matching to detect eye regions. With the help of PERCLOS establishes a driver’s personal fatigue model, system continuously monitors the driver’s state and system alerts the driver is fatigue, no.of black pixels less than a threshold means eyes are closed. Since each person has a unique eye blinking frequency and speed, a personal fatigue model is developed.Wanghua Deng and Ruoxue [6] uses yawning, blinking and eye closure duration for drowsiness detection. This system uses Multiple Convolutional Neural Networks (CNN) - KCF for face tracking, histogram equalization to improve image frame brightness, CNN for drowsiness classification, deploy models on cloud and use automobile camera.

Sukrit Mehta’s [7] model captures the driver's face in each frame by performing image processing techniques. The model calculates Eye Aspect Ratio and Eye Closure Ratio to detect driver’s sleepiness based on adaptive thresholding and has used various machine learning algorithms so as to test the efficacy and results indicated that this model can achieve 84% accuracy using random forest classifier. Another approach Kwang-Ju Kim[8] proposes is the fusion of IR image facial features and EEG features for drowsiness detection. Headphone type EEG sensors were used to decrease driver's discomfort. Rahim Soleymanpour [9] proposes Neuro Bio Monitor for sleepiness detection and it achieved an accuracy of 78.79% and 95 % detection in comparison with Karolinska Sleepiness Scale[10] and more than 70% correlation with PERCLOS[11]. Some of the mostly used machine learning algorithms for drowsiness detection are Support Vector Machine [12,13], K-Nearest Neighbor[14,15] and Random forest[16].

3. Methodology

a.Dataset

We have used a custom made dataset named as "Sleepiness Detection Dataset" that contains cropped grayscale eye images.This dataset was created by combining various dataset related to driver drowsiness available from various web sources. This dataset contains eye images with/without spectacles and reflections. 80% of dataset is used for training and 20% for validation training data out of 90,456 images.For testing our CNN model we have used 3292 images that contains with glasses and non-glasses images, different reflections, open close and close eye and so on.

b.Proposed Model

The proposed system uses a web camera to capture video of the driver's face and we convert this into frames and each frame is processed sequentially.Our proposed model consists of 4 main stages, as shown in figure (2) . First, the face region is detected and separated from the image that has been captured. Second, the two eyes are detected from the face image. Third, the mouth is detected from the face region. These two undertakings work together with one another and at that point eye state and yawn detections are applied

to the extracted mouth and eye. Finally, the outcome is combined and a decision is drawn regarding the drowsiness of the driver and if drowsy state is detected, an alarm is sent to the driver. This process is repeated in subsequent frames also. In the coming subsections, we clearly explain each stage in detail.

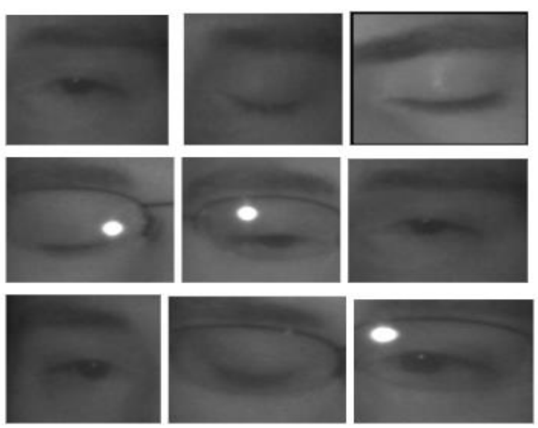


Figure 1: Cropped eye images

In a driver sleepiness detection model fusion can be applied at three stages like feature extraction , similarity checking stage, and the decision stage. In our proposed model, we have applied fusion in the decision stage.
Our model considers three levels of drowsiness :

- State 0 : In this state the eye closure is only detected and no sign of yawning.
- State 1: In this state yawning is detected, and the blink frequency is increased.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

❖ State 2: In this state eyes are mostly closed and yawning duration is increased.

i)Detection and Extraction of Face, Eyes and Mouth : We used dlib's HOG based detector and loaded facial landmark predictor in order to get the facial landmarks which uses a HOG feature descriptor with a linear SVM machine learning algorithm to detect faces. Histogram of Oriented Gradients is basically a feature descriptor most commonly used in image processing and computer vision primarily for the purpose of object detection. HOG divides the image into small connected cells and computes histogram for each cell and together forms a feature vector. Left eye, right eye and mouth are extracted using face_utils of the imutils library that contains encoded mapping of facial landmarks. Dlib implements a facial landmark detector which produces 68 (x, y)-coordinates mapping to specific facial structures and these 68 point mappings are derived by training a predictor of shape on the labeled dataset iBUG 300-W.

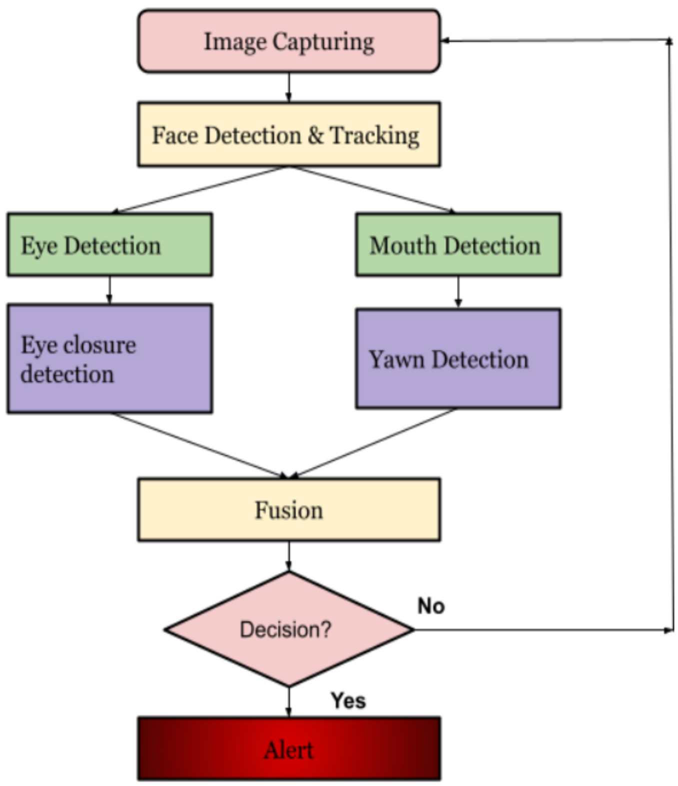


Figure 2: Main Stages of Proposed Model

ii)Eye State detection using CNN model : In our model, once the eyes are extracted it is converted into grayscale , resized to 30 x 30 and fed to CNN classifier for predicting the eye state i.e., open or closed. The data is normalized in the range 0 to 1 for better convergence. The model predicts 4 classes where 0 represents "open_left", 1 represents " open_right ", 2

represents ``close _ left " and 3 represents " close _ right". These output predictions are compared using the if else statement if any of the output predictions is 1 or 0 then eye state is OPEN else CLOSE.A CNN can be an exceptional sort of profound neural organization that performs amazingly well for picture characterizations and it comprises an info layer, shrouded layer and yield layer which can have numerous quantities of layers.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

Convolution activities can be performed by utilizing a channel that performs 2D framework duplication on the layer and channel. The CNN model architecture of the proposed model is shown in figure (3) . The model consists of 2 convolution layers with kernel of size 3 x 3 and Max pooling , flatten and two dense layers. Relu function is used in other layers except the output layer which uses softmax function. The number of nodes in the first convolution layer is 32 and second is 64. Flatten layer acts as a connection between the convolution and dense layers. The node count in the output layer is 10. The model is trained with 'Adam' as an optimizer and 'sparse categorical cross entropy as loss function. We have used 15 epochs to train and validate our model.

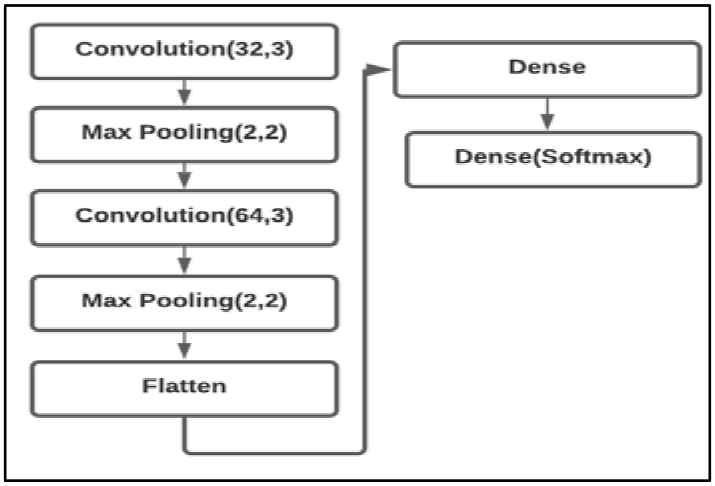


Figure 3:Proposed CNN model architecture

iii) Yawn detection using Mouth Aspect Ratio: Extracted mouth region is passed for yawn detection and this is done using Mouth Aspect Ratio (MAR). In order to detect yawning i.e., whether mouth open or close we calculate MAR. The formula to calculate MAR is mentioned in equation (1) and we have used the Euclidean distance metric to calculate the distance between the landmark points and formula is given in equation (2).

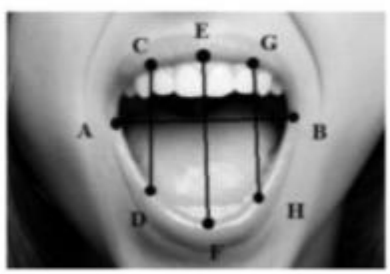


Figure 4: Coordinate points

$MAR = |EF| \div |AB|$ eq-(1)

$d = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$ eq-(2)

For yawn detection we have used the Convex Hull algorithm for obtaining the set of pixels included in the smallest convex polygon surrounding all the pixels in the input. The yawn can be recognized based on the angle proportion of the mouth that has been extracted and it is larger than a particular

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

specified limit. The edge is chosen tentatively equivalent to .65 and to inorder to make this stage more accurate, the mouth region is verified whether the map has a hole like structure and if this kind of a structure is obtained by the Convex Hull algorithm then only yawning state is detected. As mentioned
earlier the Euclidean distance between the upper and lower lip is computed and compared with the threshold value to determine the driver state i.e., awake or asleep. One important consideration is that while detecting the driver's mouth region, the distance between the person and camera has to be taken care of and the threshold value has to be selected accordingly. The space between the upper and the lower lip will decrease in the image when the person moves away from the camera . Hence, a required threshold value is selected with reference to the distance between the camera and person.

iv)Decision Making Stage: In this stage a decision on the driver's sleepiness state is made and an alert is sent to the driver if : The eye closure and

yawning both are detected simultaneously, Consecutive yawning is detected for successive frames or Consecutive eye closure is detected for successive frames. Hence if frequent yawning or eye closure is detected for 30 consecutive frames with frame rate of 25 frames per second then an alert will be sent to the driver. Complete model architecture of the proposed model is depicted in figure(5).

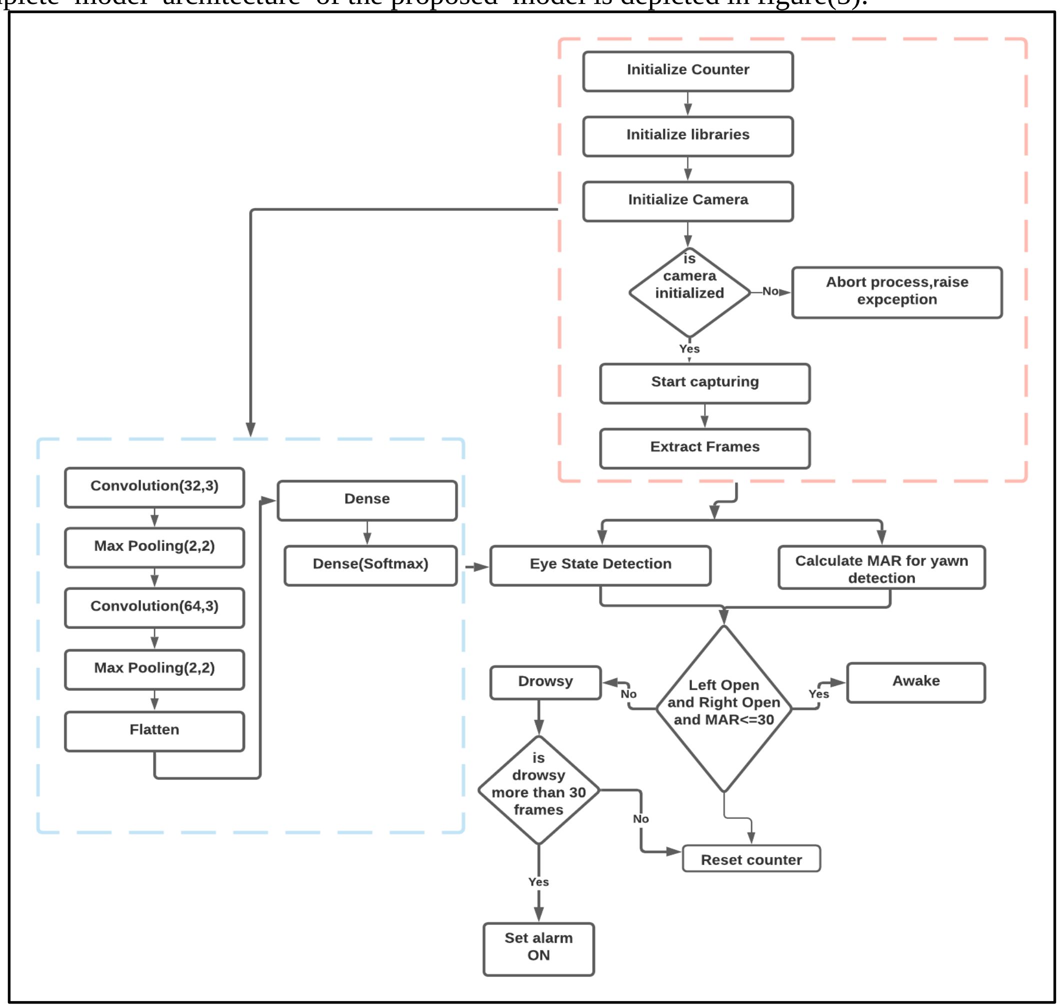
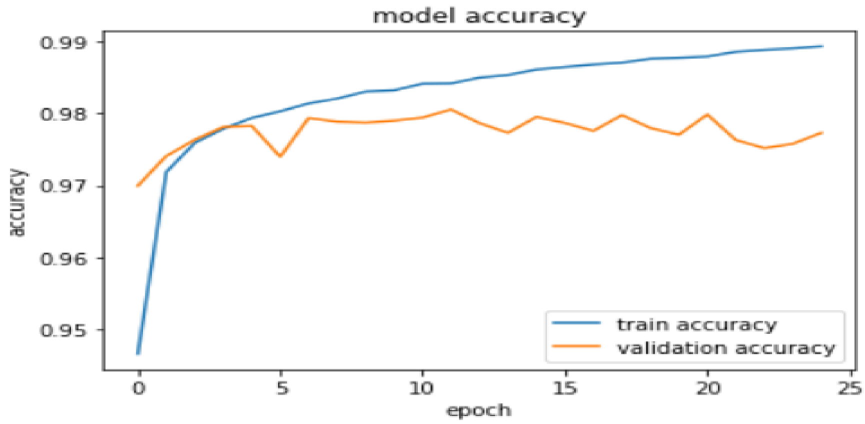


Figure 5: Proposed Model architecture

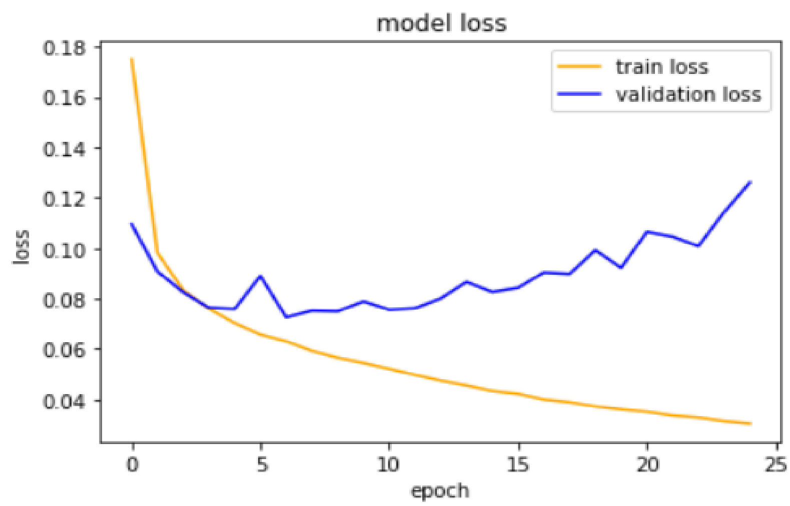
4. Experimental Results

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

This section presents the performance evaluation for the proposed approach by performing an empirical analysis of obtained results. The system collects the real-time data of the drivers and then determines drowsiness of the drivers based on the values that are computed from the images captured. The system alerts the driver if drowsiness is detected. Figure [6a] shows training accuracy increases as the epoch increases and in figure[6b] training loss decreases as epoch increases.



{a}



{b}

Figure 6 : Model Accuracy and Loss

In the last epoch that is 25th , training accuracy is 98.93% and validation accuracy is approximately 97%.To test our CNN model, we have used a total 3292 images consisting of images of various lighting conditions, reflection, with/without glasses and achieved test accuracy of 97.65%. Our model was compared with other mostly used machine learning algorithms in drowsiness detection so as to compare performance and it was found that our model out performs other models and this can be understood from the table provided below.

Algorithm	Accuracy	Precision	Recall	F1 Score
Random Forest	95.56%	0.956	0.956	0.956
MLP	95.79%	0.958	0.958	0.958

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2248-2256
https://publishoa.com
ISSN: 1309-3452

Proposed	97.65%	0.976	0.977	0.976

Table 1 Comparative Analysis with other Algorithms

5.Conclusion

The importance of this research paper stems from the need for an early drowsiness detection so that we can have a decrease in the count of road accidents due to driver drowsiness or fatigue. Drowsiness leads to late response which causes accidents. So we proposed a model that detects the driver's drowsiness by monitoring the behavioral characteristics of the driver such as eye state and yawning. If the driver is detected drowsy then the model alerts the driver that he/she is drowsy and has to take rest. Both eye closure and yawn are detected to increase the model performance and accuracy. We used CNN for eye state prediction and MAR for yawn detection. Outcomes demonstrate high effectiveness of the model. We assume a threshold distance between the lower and upper lip for the subject's yawn. The proposed framework works well at different lighting conditions and also if the driver wears spectacles. The framework has an accuracy of 98.75% for CNN classifier.

References

1. Md. Tanvir Ahammed Dipu , Syeda Sumbul Hossain, Yeasir Arafat, Fatama Binta Rafiq, "Real-time Driver Drowsiness Detection using Deep Learning",(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 7, 2021.
2. Jabbar, R., Shinoy, M., Kharbeche, M., Al-Khalifa, K., Krichen, M., & Barkaoui, K. (2020, February). "Driver drowsiness detection model using convolutional neural networks techniques for android application". In 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT) (pp. 237-242). IEEE.
3. Jonathan Flores-Monroy, Mariko Nakano-Miyatake,Gabriel Sanchez-Perez,Hector Perez-Meana, "Visual-based Real Time Driver Drowsiness Detection System Using CNN", 2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 10-12, 2021.
4. R. Ayachi, M. Afif, Y. Said and A. B. Abdelali, "Drivers fatigue detection using EfficientDet in Advanced Driver Assistance System," in Proc. Int. Multi-conference on Systems, Signal & Devices, 2021, pp. 738-742.
5. Jun-Juh Yan, Hang-Hong Kuo , Ying-Fan Lin , Teh-Lu Liao, "*Real-time Driver Drowsiness Detection System Based on PERCLOS and Grayscale Image Processing*",978-1-5090-3071-2/16 $31.00 © 2016 IEEE DOI 10.1109/IS3C.2016.72
6. Wanghua Deng and Ruoxue WU,"*Real-Time Driver-Drowsiness Detection System Using Facial Features"* ,Digital Object Identifier 10.1109/ACCESS.2019.2936663.
7. Sukrit Mehta, Sharad Dadhich, Sahil Gumber, Arpita Jadhav Bhatt, "Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio",© 2019SUSCOM. Hosted by Elsevier SSRN. All rights reserved.
8. Kwang-Ju Kim,Kil-Taek Lim,Jang woon Baek,Miyoung Shin, "Low-Cost Real-time Driver Drowsiness Detection based on Convergence of IR Images and EEG Signals", 021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) | 978-1-7281-7638-3/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICAIIC51459.2021.9415193.
9. Rahim Soleymanpour, Hossein Hamidi Shishavan, Jae-Sang Heo, and Insoo Kim, "Novel Driver's Drowsiness Detection System and its Evaluation in a Driving Simulator Environment", 021 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 978-1-6654-4207-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMC52423.2021.96592.
10. M., Ingre, T. Åkerstedt, B. Peters, A. Anund, G. Kecklund, "Subjective sleepiness, simulated driving performance and blink duration: examining individual differences," Journal of sleep research, 15(1), 2006, pp. 47-53.

11. Raja Mohana SP, Manu Vidhya S, Reshma D, "A Real-time Fatigue Detection System using MultiTask Cascaded CNN Mode", 021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT) | 978-1-6654-2306-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSNT51715.2021.950962.

12. Salma Anber, Wafaa Alsaggaf and Wafaa Shalash, "A Hybrid Driver Fatigue and Distraction Detection Model Using AlexNet Based on Facial Features ", Electronics 2022, 11, 285. https://doi.org/10.3390/electronics 11020285.

13. Zhenlong Li, Qingzhou Zhang and Xiaohua Zhao, "Performance analysis of K-nearest neighbor, support vector machine, and artificial neural network classifiers for driver drowsiness detection with different road geometries",International Journal of Distributed Sensor Networks 2017, Vol. 13(9) The Author(s) 2017 DOI: 10.1177/1550147717733391 journals.sagepub.com/home/ijdsn.

14. Prima Dewi Purnamasari,Pratiwi Yustiana,Anak Agung Putri Ratna,Dodi Sudiana , "Mobile EEG Based Drowsiness Detection using KNearest Neighbor ",IEEE Xplore.

15. Amir Jalilifard and Ednaldo Brigante Pizzolato, "An Efficient K–NN Approach for Automatic Drowsiness Detection Using Single–Channel EEG Recording",978-1-4577-0220-4/16/$31.00 ©2016 IEEE.

16. Monica Abad, James Carlisle Genavia, Jaybriel Lincon Somcio, Larry Vea, "An Innovative Approach on Driver's Drowsiness Detection through Facial Expressions using Decision Tree Algorithms", 0571 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) | 978-1-6654-0690-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/UEMCON53757