

Enhanced Driver Drowsiness Detection using Deep Learning

Dipender Singh^{1*} and Avtar Singh²

¹Department of Computer Science and Engineering, Dr B R Ambedkar National Institute of Technology, Jalandhar, 144011, Punjab, India

²Department of Computer Science and Engineering, Dr B R Ambedkar National Institute of Technology, Jalandhar, 144011, Punjab, India

Abstract- The primary reason for road accidents is drowsiness reported by National Highway Traffic Safety Administration (NHTSA). To overcome this issue, researchers have proposed and implemented various methods based on driver behaviour and vehicle movements. Vehicle-based methods often rely on a set of predetermined parameters to detect drowsiness, such as changes in steering wheel angle or lane deviation. However, these parameters may not always accurately reflect a driver's level of alertness. Therefore, it is essential to develop an effective approach for driver drowsiness detection. Deep learning techniques such as convolutional neural networks (CNN) are structured solutions to detect drowsiness based on drivers' facial features. The proposed approach based on CNN focuses on the eyes and mouth region using the nose as a central point. CNN is operated with rectified linear activation function (ReLU) which gives 94.95% accuracy as compared to existing methods even in different situations namely low light, different angles, and transparent glasses.

Keywords- Convolutional Neural Network (CNN), Electrooculography (EOG), and Rectified linear activation function (ReLU).

1 INTRODUCTION

One of the primary causes of road accidents is driver drowsiness [1]. Regularly driving for very long period, drivers are certainly exhausted, which leads to drivers being tired and sleepy. After many researches, it was found that the main reason for road accidents is drowsiness. Different countries have distinct reports. From all reports, it is found that driver fatigue is the main cause of accidents. Developing a method that determine driver fatigue is a challenge. According to the report of "Ministry of Road Transport and Highways" there were 16231 accidents in India in 2022. Due to the increasing number of vehicles, the number of accidents is increasing day by day. Numerous vehicles are driving at night, including heavy equipment trucks. Drive these kinds of vehicles continuously for such a long period of time become more prone to these drastic conditions.

During last two decades various methods including GPS systems that detects when a driver begins to drift out of their lane due to inattention. Some other techniques used electrocardiographic equipment, which measure the driver's heart rate to determine the driver's sleepiness. One of the main goals of intelligent transportation systems (ITS) is to increase people's well-being and minimize the number of accidents. One of the main factors in accidents, especially in gridlock roads is the drowsiness and monotony of drivers. The ability to make decisions and the awareness to drive a vehicle is reduced due to drowsiness. According to some research, drivers feel drowsy after short period of driving, most particularly in the early afternoon, after and late at night, when the driver's tiredness is high and the state of mind is far from normal. Moreover, hypnotic substances and alcohol use cause the driver to lose awareness. In other Nations, completely different types of accident stats were reported, which were caused by driver's distraction. The most common cause of two hundred accidents is lack of concern and the transitory condition of the driver. Studies show that the use of a drowsiness detection system significantly reduces the number of accidents by at least 20% [1]. In order to detect the driver's drowsiness, a series of facial images are captured in which eye blink and head movements are monitored. Facial pattern exploration can be a fashionable area of analysis that has various applications such as facial recognition, security systems for people authentication, and virtual tools.

Proposed approach focuses on the localization of the driver's eyes and mouth which is achieved by capturing a complete representation of the face from the data set that includes 1400 images taken from Kaggle and then detecting the position of the eyes and mouth using a Convolution neural network. After detecting the desired regions, the method detects whether the driver is drowsy or not based on eye and mouth movements. Develop a

* Corresponding author : dipenders.cs.21@nitj.ac.in

CNN-based system that can accurately detect driver drowsiness in the data set which includes 1400 images of different people with different conditions. The system should be able to identify common signs of drowsiness, such as heavy eyelids, yawning, and nodding off, and alert the driver to take a break. The need is to design a system that works under various lighting conditions that must be robust enough to handle variations in head movements and facial expressions. The paper is divided into subparts which include the previous works, limitations of the present system, methodology and experimental results.

2 Previous Works

This section followed the various recent works to achieve the better results in drowsiness detection.

Sukrit et al. in [2] have used the most commonly used technique which is the Eye Aspect Ratio (EAR) and Eye Closure Ratio (ECR). Same accuracy for people wearing spectacles. Useful in conditions when drivers drive for longer distances. Adaptive thresholding varies greatly from person to person. So, In [2] Deng and Wu in [3] have used video Images that were used to detect Blinking, yawning and duration of eye closure. They drafted an identification technique for facial regions which is based on the Golden Ratio of 68 main points. There were not any publicly available image-based driver sleepiness recognition datasets. algorithm works with an accuracy reaching 95% when the Euclidean distance is within 20 pixels. Average accuracy regardless of environmental conditions is approximately 92%. The limitation is that it requires a high-performance machine with heavy processors and good RAM.

An artificial Neural Network is used in [4]. Focusing more on EEG (Electroencephalography) and EOG signals measurement and image classification. High complexity using devices Signals. Works effectively in bad conditions. The problem is Collecting data from signaling devices is difficult to handle these complex devices. In [5] an important ensemble machine learning is used based on hybrid sensing. Making observations based on driving performance using counterfeit and driver observation systems. It has a high accuracy of 82.4% using driving. The number of candidates who participated was not sufficient and all were males in their 20s. Anchan and Saritha have used Image Processing Algorithms to detect the face, position of the head, and Blinking of the eye[6]. This system is available as an android application. So, there is no extra cost. It has a low accuracy of 81% and thus is unreliable.

Yan et al. [7] have used face detection systems to detect dullness and lethargy using the Percentage of eye closure, eyelid separation and eye shutdown rate. The eyelid separation and percentage of eye closure are used to find how much a person is already exhausted and the eye separation is used for the detection of diversions. This method focuses on three different parts of the face, thus providing more options for detecting fatigue and distraction. The limitations of this method are unable to measure fatigue and dizziness. Therefore, object evaluation is not possible using this directly. PERCLOS (Percentage of Eye Closure) and Grayscale Image Processing are used for drowsiness detection by kuo and Fan-Lin. It uses grey-scale images which greatly reduces the memory space required to store images compared to RGB images. Accuracy is high, more than 90% when the driver's eyes are open, and the proportion is greater than 50%. The system requires additional calculation steps.

The publication[8] suggests using blink patterns and horizontally symmetrical eye features to detect sleepiness. Detects eye blinks with over 90% accuracy and 1% false positives. It has an advantage over systems that use statistical information from past frames, because this system is not dependent on changes in head position, it works on the same time frame. It only provides results for blink detection as there is no common database to compare its result for sleepiness. Its performance is affected by the presence of glasses. Electrooculography is proposed by Parisa et al. in [9]. It is based on detecting eye blinks, and horizontal symmetry features of eyes. For saccade detection, the proposed algorithm has false drowsy rate less than 5% and true positive greater than 80%. The eye blink discernment using this suggested algorithm will perform better than the normal median filter by at least 20%. In this system, 3 means clustering is required to prevent confusion of driver's reduced eyelash drops including any eye movements which suspect drowsiness.

In [10] has used various sensors including heart rate, breathing and visual signal for the act. It has an accuracy of 93% and it focuses on four different types of distractions. It was an analysis for detecting different types of distractions but the driver was not alerted when he/she got distracted. Schwarz et al [11] have used NADS (National Advanced Driving simulator). The models are useful in detecting low levels also. Eyelash Drops are predictive microsleeps.[12] Discusses drowsiness of the operator using Image Processing which includes sensing using driver operation, physiological factors and vehicle response. This system can handle various conditions such as light changes, shadows, reflections, etc. The system also works regardless of the texture and color of the face. During intensive use, water vapor can get on the used sensors, which leads to a reduction in the utilization capacity. In [13] Proposed an Eye Aspect Ratio to detect features of the user from various angles to drowsiness detection. It shows good results in Rainy Weather, Wear glasses (transparent) with bright light and dim light. The method fails to detect drowsiness under some conditions like Wear sunglasses and Night drive.

Table 1. Comparative analysis of various techniques

Ref	Problem	Need	Technique Used	Parameters	Solution	Environment	Future Direction
[3]	Purposed a DriCare System that detects the yawning, eye closure duration and blinking with the help of the video images.	To detect the driver drowsiness with more tracking accuracy.	Optimizes the algorithm KCF with the help of purposed Multi convolutional Neural Network.	Detect the eyes and mouth movements with the help of the Dlib Library	Developed a facial region detection technique based on 68 essential components with 92% accuracy.	Data Set is generated with the help of the different team members to collect the dataset from the vehicle camera in same conditions.	Data set should be based on different public in different conditions.
[2]	Develop a system(application) install on the driver car to track the eyes movement for driver drowsiness.	To provide the more accuracy in different conditions.	Eye Closure Ratio and Eye Aspect Ratio is used to detect the driver drowsiness.	Detection of the face with the help of linear SVM and histogram	System works with almost same accuracy with spectacles and useful in conditions when drivers drive for long distances.	Method based on the real-time detection of the driver's face with the help of the application installed on the system.	Adaptive thresholding varies greatly from person to person. Thresholding based on the other criteria rather than sleep count
[5]	Detection system based on the factors: vehicle, physiological indicators and behavioural using hybrid sensing.	Classified the driver's alert state mainly slightly drowsy state for early drowsy detection.	For the Physiological signals: electroencephalogram and electrocardiogram, driving simulators for behavioural indices.	Combination of the three techniques: ECG and EEG for the physiological signals, simulators for driver behaviour and vehicle classification using sensors.	System works on every field using combination of the technologies with accuracy of 82.4%.	Driver is covered with sensors on body and surrounding.	System is expensive using these methods. System must be adaptable for everyone.
[4]	Three types of methods introduced for driver drowsiness: EEG, EOG and image analysis. Artificial Neural networks used to analyse the eyes state.	Detection of driver drowsiness using different methods for better results.	Sensors of EEG and EOG fixed for the brain and eyes activity. ANN is used for analysis of the images.	Detection of brain, eyes activity done using sensors fixed. Images analyse using ANN.	Work efficiency of system is good in bad conditions.	Sensors and camera used for physiological signals and driver's images.	Collecting data and handling EEG and EOG signaling devices is difficult.
[14]	Driver's eyes and mouth location find using their geometric features which helps in drowsiness detection.	Detection of drowsy using state of eyes and mouth for better accuracy.	Two methods are introduced: Percentage and Time That Mouth is Open (PTMIO), Percentage and Time that Eyelids cover the Pupils (PATECP).	Detect eyes and mouth location.	AdaBoost algorithm used for face detection including edge and line features which leads to 84% accuracy.	Image is captured using camera and carrying out for further process.	Accuracy of method is less in bad conditions and different people.
[11]	Detection of driver drowsiness using sensors for monitoring driver movements.	For analysing level of drowsiness from low level to high level.	NADS (National Advanced Driving simulator) and signals from the vehicle's sensors integrated with the driver monitoring system.	Detection of vehicle-based activities like handle movement, phone use and distractions.	Models performed well in identifying mild tiredness from moderate to severe drowsiness.	Sensors fixed in different locations of vehicle.	Blinks have been shown to be predictive of microsleeps.

Shen et al. in [14] cites a user's sleepiness based on multiple facial features using the percentage and time of the eyelids covering the pupil, the percentage of time the mouth is visible to the camera. Detects sleepiness using eyes and mouth. The disadvantage is that the extraction of targets is difficult in some cases and sometimes the behavior of the target is not detectable. Khosro et al. [15] suggested computing the drowsiness of operators using Haar-like features and AdaBoost classifiers. Local Binary Pattern is employed to fetch characteristics of eye features. Mohit Dua et al. [16] suggested combination of different deep learning models like Rest Net and Alex net using RGB videos as input and detect drowsiness. Model uses GPU for better accuracy to detect facial expressions and hand gestures. Model is divided into subparts and every deep learning model has their own functioning respectively.

Some disadvantages of driver drowsiness detection systems at present include high false alarm rates, difficulty in detecting microsleeps, and the need for calibration or individual customization. Additionally, some systems can be intrusive or uncomfortable for the driver, and their effectiveness can be affected by environmental factors. Present drowsiness detection systems use complex approaches and costly equipment, such as an Electrooculogram (EOG) which monitored the movements of muscles acting on the eyes. Some methods work on face detection using 68 facial points which store coordinates in dynamic storage. But these methods are not able to find the accurate desired regions if the full face is not in the frame as all coordinates are correlated to each other. In recent years most of the methods follow the Convolution neural networks (CNN) which give more accuracy but fail in different conditions like different angles, low light, and transparent glasses. The main reason for failure is the Working of CNN as it required different angles (0, +200, -200) to analyze the object. Angle problem in all driver-based techniques remain same because every technique just focuses on increasing the number of hidden layers which causes the loss of the lot of data.

3 Methodology

Main motive of this paper to provide a method which gives more accuracy in different conditions. For this, method is divided into subparts that work efficiently. The proposed method implement on dataset includes 1400 images taken from Kaggle which includes different conditions. The following steps are used to solve the issues of driver drowsiness as shown in figure 1.

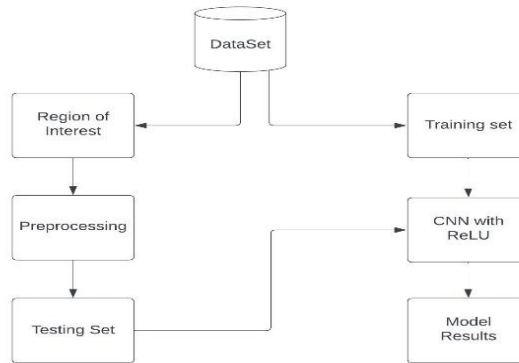


Fig. 1. Steps of Methodology.

Step1:

First step of methodology after loading dataset in machine is testing and training. Most common split ratio of training and testing is 80% and 20% respectively for dataset of images more than 1000. But chances of overfitting are very high to achieve more accuracy. To overcome this situation, k-fold cross validation method is used which gives 1078 images as training and 3022 images for testing which eliminates overfitting chances.

Step 2:

Next Step of methodology to detect eyes and mouth region for further process. As earlier mention in paper, driver looked at different angle so concerned regions sometimes lost. To overcome this problem, detection of eyes and mouth region done based on nose coordinates. Some of sample figures are mentioned below for reference how interested region is detected.

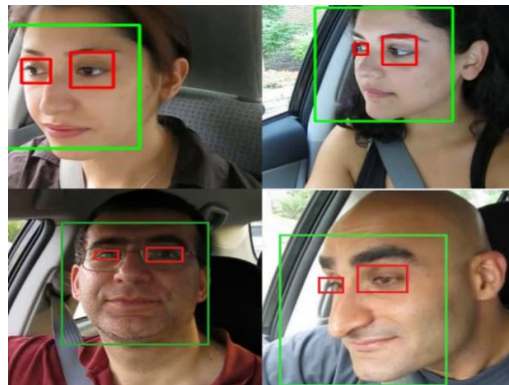


Fig. 2. Detection of Eyes Region

The region is segmented with Convolutional Neural Network (CNN) model based on pre-trained model of human eye, mouth and face. The trained CNN model uses the Images pre-processed by TensorFlow library. Method uses CNN python libraries to detect eyes and mouth region without any manual image processing. The trained model can be applied to different images with different features and values. Class feature of TensorFlow to differentiate between closed, open and other facial expressions. We use an open-source library named OpenCV for training and analysing images.

Step 3:

Model takes a trained Convolutional Neural Network (CNN) and converts the weights into simplex and maxpooling layers. The Sequential model is a higher-level neural network architecture that implements sequence learning. The model takes a sequence of zero or more random inputs and predicts a single output, which can be used as an input to subsequent layers in the network. Model fit over a training set of 30 epochs with validation data that are generated based on the output of that model. The validation steps are chosen to yield the best training accuracy, while minimizing time and memory usage. The final model is then flattened, dropped out, and trained on the given dataset.

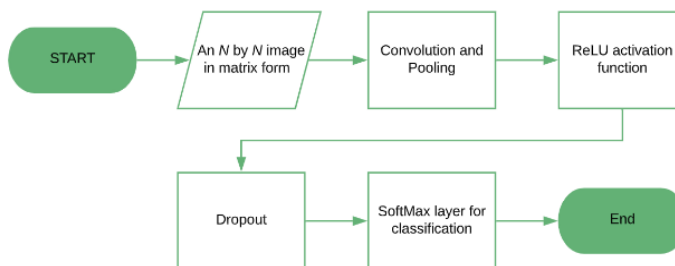


Fig. 3. Working of CNN with ReLU

This model has 3 convolutions, 4 max pooling and 1 flatten on top of the previous layers. The final output of this model is a dense vector with 64 dimensions per class. In a CNN, ReLU is usually applied after the convolutional and pooling layers. The output of the convolutional layer is passed through the ReLU activation function, which sets all negative values to zero and leaves positive values unchanged. This helps to eliminate any negative values that might arise in the output of the convolutional layer, which can hinder the performance of the network. The use of ReLU with CNNs has been shown to improve the accuracy of the final results, especially in image recognition tasks.

4 Experimental Results

In this study, the proposed approach aimed to develop a new CNN-based driver drowsiness detection system using a dataset of frontal face images. The dataset consisted of 1400 images with a balanced ratio of normal images and drowsy images. The images were preprocessed to resize them to 145x145 pixels and normalize the pixel values.

We trained and evaluated the proposed CNN model using Keras with the TensorFlow backend. The model consisted of three convolutional layers, each followed by a max pooling layer, and two fully connected (dense) layers. The proposed approach uses ReLU activation function for the convolutional layers and the SoftMax activation function for the output layer. The model was trained with a batch size of 39 and for 30 epochs, using the Adam optimizer and categorical cross-entropy loss function. CNN model was able to achieve an accuracy of 94.95%, which is better than other existing techniques for driver drowsiness detection.

Table 2. Different Models Comparison

Model	Technique	Accuracy (%)
Proposed Approach	Eyes and Mouth Region	94.95
CNN-based	Eyes Tracking	92.3
SVM-based	EEG	88.9
Random Forest	Heart Rate Variability	91.2
LSTM-based	Driving Behavior Analysis	87.5
Rule-Based	Steering Wheel Movements	82.1

The results showed that the model achieved an accuracy of 94.95%, which is significantly better than other techniques used for driver drowsiness detection. This high accuracy level demonstrates the effectiveness of the improved CNN model in detecting driver drowsiness.

The graph in figure 4 shows the accuracy of a CNN model over 30 epochs, with two labels: train accuracy and test accuracy. The train accuracy represents the accuracy of the model on the training dataset, which is the data that the model was trained on. Initially, the train accuracy is low, as the model is still learning and improving its performance. As the model is trained on more data, the train accuracy improves, reaching a plateau where it no longer improves significantly.

The test accuracy, on the other hand, represents the accuracy of the model on a separate dataset that it has not seen before. This is an important metric, as it indicates how well the model will perform on new, unseen data. Initially, the test accuracy is low, as the model has not yet learned to generalize well to new data. The train accuracy and test accuracy constantly increase as the number of epochs increases.

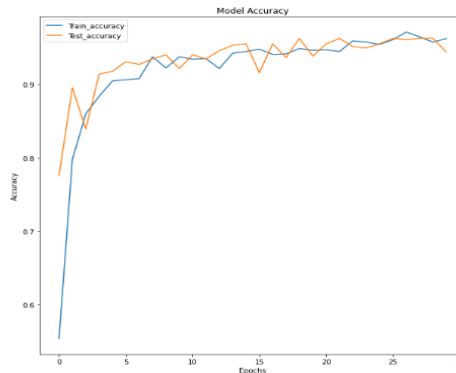


Fig. 4. Model Accuracy

Ideally, the train accuracy and test accuracy should follow a similar trend, with both increasing as the model is trained on more data.

The graph in figure 5 shows the model’s loss over 30 epochs, with two labels: train loss and test loss. It indicates how well the model is fitting the data, with a lower loss indicating better performance. Overall, both graphs provide valuable insight into the performance of the CNN model over time, allowing for adjustments to be made to improve its accuracy and generalization capabilities.

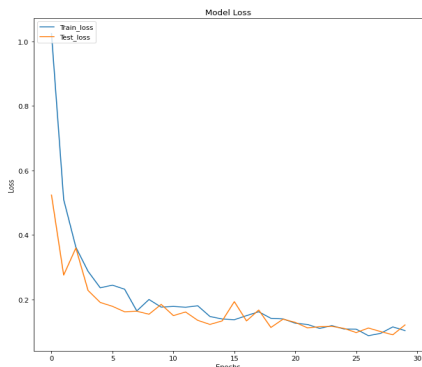


Fig. 5. Model Loss

In figure 6, the Proposed Model experimental results showed that our CNN-based driver drowsiness detection system achieved an accuracy of 94.95% in detecting driver drowsiness, with a precision of varies from 80% to 99% and recall of 82% to 98% for different factors including in dataset. The system also achieved an F1 score of 88% to 97% and an AUC of 0.95. The chances of model overfitting increases the number of epochs increases. The change in the model remains overall same after 30 epochs. Overall results of the model are better in different conditions.

	precision	recall	f1-score	support
yawn	0.95	0.82	0.88	45
no_yawn	0.80	0.98	0.88	45
Closed	0.99	0.94	0.96	216
Open	0.95	0.98	0.97	229
accuracy			0.95	535
macro avg	0.92	0.93	0.92	535
weighted avg	0.95	0.95	0.95	535

Fig. 6. Model Performance Summary

5 Conclusion

To overcome the problem of road accidents that significantly happened due to driver drowsiness. The proposed model focuses on closed and open eyes with yawning cases which are major factors in detection. The main aim is to detect drowsiness in different conditions for better accuracy. Overall, the model experimental results demonstrate the potential of using CNN for driver drowsiness detection, which can be further improved by using larger datasets and more advanced CNN architectures. The experimental results also highlight the importance of developing robust and reliable driver drowsiness detection systems to improve road safety and prevent accidents.

References

- [1] K. Kumari, “Review on Drowsy Driving: Becoming Dangerous Problem,” International Journal of Science and Research, (2014).
- [2] S. Mehta, S. Dadhich, S. Gumber, and A. J. Bhatt, “Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio Fatigue Detection Non-Intrusive Methods Driver monitoring system,” (2019).
- [3] W. Deng and R. Wu, “Real-Time Driver-Drowsiness Detection System Using Facial Features,” IEEE Access, vol. 7, pp. 118727–118738, (2019).
- [4] T. Vesselenyi, S. Moca, A. Rus, T. Mitran, and B. Tătaru, “Driver drowsiness detection using ANN image processing,” in IOP Conference Series: Materials Science and Engineering, vol. 252, no. 1, (2017).

- [5] J. Gwak, A. Hirao, and M. Shino, "An investigation of early detection of driver drowsiness using ensemble machine learning based on hybrid sensing," *Applied Sciences*, vol. 10, no. 8, (2020).
- [6] S. L.R. and S. Anchan, "Human Drowsiness Detection System," *International Journal of Engineering and Advanced Technology*, vol. 8, no. , pp. 316–319, Apr. (2019).
- [7] J. J. Yan, H. H. Kuo, Y. F. Lin, and T. L. Liao, "Real-time driver drowsiness detection system based on PERCLOS and grayscale image processing," *IEEE International Symposium on Computer, Consumer and Control, IS3C* ,pp. 243–246,(2016).
- [8] T. Danisman, I. M. Bilasco, C. Djeraba, and N. Ihaddadene, "Drowsy driver detection system using eye blink patterns," in 2010 International Conference on Machine and Web Intelligence, *ICMWI 2010 - Proceedings*, pp. 230–233,(2010).
- [9] P. Ebrahim, W. Stolzmann, and B. Yang, "Eye movement detection for assessing driver drowsiness by electrooculography," in *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*,pp. 4142–4148,(2013).
- [10] M. Gjoreski, M. Z. Gams, M. Luštrek, P. Genc, J. U. Garbas, and T. Hassan, "Machine Learning and End-to-End Deep Learning for Monitoring Driver Distractions from Physiological and Visual Signals," *IEEE Access*, vol. 8, pp. 70590–70603, (2020).
- [11] C. Schwarz, J. Gaspar, T. Miller, and R. Yousefian, "The detection of drowsiness using a driver monitoring system," *Traffic Inj Prev*, vol. 20, no. sup1, pp. S157–S161, (2019).
- [12] S. Himani Parmar, M. Jajal, and Y. Priyanka Brijbhan Electronics, "Drowsy Driver Warning System Using Image Processing | Drowsy Driver Warning System Using Image Processing," *IJEDR*, (2019).
- [13] C. B. S. Maior, M. J. das C. Moura, J. M. M. Santana, and I. D. Lins, "Real-time classification for autonomous drowsiness detection using eye aspect ratio," *Expert Syst Appl*, vol. 158, (2020).
- [14] P. Wang and L. Shen, "A method of detecting driver drowsiness state based on multi-features of face," in 2012 5th International Congress on Image and Signal Processing, *CISP 2012*, pp. 1171–1175,(2012).
- [15] M. Sabet, R. A. Zoroofi, K. Sadeghniaat-Haghighi, and M. Sabbaghian, "A new system for driver drowsiness and distraction detection," in *ICEE 2012 - 20th Iranian Conference on Electrical Engineering*, pp. 1247–1251,(2012).
- [16] M. Dua, Shakshi, R. Singla, S. Raj, and A. Jangra, "Deep CNN models-based ensemble approach to driver drowsiness detection," *Neural Comput Appl*, vol. 33, no. 8, pp. 3155–3168, (2021).
- [17] S. Liu, Y. Li, C. Wen, and S. Wang, "A Novel Driver Drowsiness Detection Method Based on Heart Rate and Its Variability," *IEEE Access*, vol. 9, pp. 114103-114114,(2021).
- [18] F. Li, X. Zhao, M. Liu, and H. Han, "Driver drowsiness detectionbased on multi-task learning and EEG signals," *Computer Methods and Programs in Biomedicine*, vol. 205, (2022).
- [19] S. Anwar, K. Khan, A. Rehman, and M. Baig, "A novel driver drowsi- ness detection system using non-contact vital signs and machine learningtechniques," *Biomedical Signal Processing and Control*, vol. 68, (2021).
- [20] C. Liu, J. Lu, Y. Zhang, and Y. Yang, "Driver drowsiness detection using multi-task learning and a convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 1748-1758,(2022).
- [21] H. He, Y. Zhou, X. Zhang, and J. Zhang, "Driver drowsiness detection using multi-view convolutional neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 1248-1258,(2022).
- [22] H. Li, Z. Li, Y. Li, and Y. Li, "A novel driver drowsiness detection method based on deep learning and physiological signals," *Measurement*, vol. 190,(2022).
- [23] N. K. Ibraheem and S. S. Yahya, "Real-time driver drowsiness detectionsystem using machine learning algorithms," *SN Computer Science*, vol. 2, no. 3,(2021).
- [24] A. Khan, A. Qayyum, and H. Bokhari, "Driver drowsiness detection system using deep learning and handcrafted features," *IEEE Access*,vol. 9, pp. 64416–64427,(2021).
- [25] A. Sharma and R. Pandey, "Efficient driver drowsiness detection using transfer learning with deep neural networks," *Computers Electrical Engineering*, vol. 91,(2021).
- [26] S. Reddy, S. K. Arora, and S. K. Rautaray, "Driver Drowsiness Detectionusing Deep Learning and Effective Fusion of Multi-modal Data," *PatternRecognition Letters*, vol. 154, pp. 11-20,(2022).