ISSN: 2347-6540

International Journal of Latest Research In Engineering and Computing (IJLREC)
Volume 12, Issue 1 : Page No. 1-10
www.ijlrec.com

REAL TIME DRIVER DROWSINESS DETECTION USING YOLOV5

Aditya Yadav[1], Satender Yadav[2], Simant Gautam[3], Jugnu[4] Mukul Motla[5] Prof. Ashutosh Pradhan[6]

Student, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, adityayadav7838@gmail.com[1]

Student, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, satendar10nov@gmail.com[2]

Student, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, simant8505@gmail.com[3]

Student, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, jgsharma270@gmail.com[4]

Student, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, mukul187@gmail.com[5]

Professor, MCA, RD Engineering College, Ghaziabad, Uttar Pradesh, India, ashutoshpra@gmail.com[6]

Abstract: Due to their diverse driving features, human drivers have different driving skills, expertise, and attitudes. In order to prevent traffic accidents, it is crucial to develop an efficient sleepiness detection algorithm. Driver fatigue has been a major problem jeopardizing road safety. The challenge of identifying abnormal human driving behaviour has been tackled by a number of research projects using computer vision techniques to analyze the driver's frontal face and the dynamics of the vehicle. Nevertheless, complex characteristics of driving behaviour are not captured by the traditional approaches. But since the development of deep learning architectures, a great deal of research has also been done to use neural network algorithms to assess and identify tiredness in drivers. This research presents a novel framework for driver sleepiness recognition based on YoloV5 architectures. In order to extract Region of Interest (ROI), a unique YoloV5 pre- trained architecture is proposed for face extraction. This work presents binary image classification, which are trained and verified using the public dataset and custom dataset, in response to the shortcomings of earlier models. The model's training and validation accuracies were 96.2% and 97.4%, respectively. It achieves 95.5% accuracy. Experiments carried out indicate that our approach has a great deal of potential for real-world use in intelligent transportation systems.

Keywords: YoloV5, face detection, image categorization, drowsiness detection

I. INTRODUCTION

Drunk driving causes catastrophic injuries to many drivers and pedestrians since Advanced Driver Assistance Systems (ADAS) do not have sleepiness detecting systems. Roughly 40% of all highway fatalities and injuries are caused by sleepy or fatigued drivers, according to the Central Road Research Institute (CRRI). The National Highway Traffic Safety Administration (NHTSA) [1] reported that around 1,00,000 incidents involving sleep-deprived drivers occur annually, with 2000 fatalities and 70,000 injuries. Furthermore, individual vehicle run-offroad crashes—in which a driver loses control of their vehicle and finally leaves their lane or collides with the rear of the car ahead—account for about 80% of sleepiness-related auto accidents. As a result, sleepy driving is a major and latent risk factor for auto accidents [2].

Therefore, creating a methodical sleepiness detection algorithm is essential to lowering the number of traffic accidents.

In recent years, detection of drowsy driving has become a major research field. Recent research has divided drowsiness detection methods into three groups: facial analysis, vehicle- based measures, and physiological measures [3], [4]. First of all, physical conditions change as a driver gets weary, therefore physiological measurements depend on bodily components like heart rate, body temperature, pulse rate, and so on [5]. In general, ECG [6], EEG [7], EMG, and EOG [8] (driver) are the physiological signals that are most frequently employed to assess a person's physical condition. The main disadvantage of physiological methods is that driver comfort during sensor wearing must be guaranteed.[9] Second, vehicle-based measures detect fatigue by observing driving behaviours such irregular braking, abrupt speed changes, and movement of the steering wheel. Various car parts have sensors installed in order to measure driving performance and identify driving patterns, which helps in the detection of fatigue. The main drawback of using vehicle-based solutions is that adverse weather, bad road conditions, heavy medicine, and other circumstances can alter the behaviour of the vehicle [10]. Lastly, analysing facial expressions with computer vision (CV) and machine learning and movements, behavioural measurements or face analysis identify tiredness [11].

In recent years, a variety of behavioral strategies have been developed to recognize fatigue. Various techniques were used in several studies to identify signs of tiredness and facial expressions. Sleepiness detection can be achieved with the use of Canny edge detection [3], ViolaJones (Haarcascade) [12], and neural network techniques including CNN [13], ANN [14], Naive Bayes classifier [12], and GAN's [15]. Additionally, these techniques are used to identify face expressions. In this research, we provided a behavioral approach architecture that starts with face identification and ends with drowsiness detection using Yolo V5. The primary contributions of this study are as follows:

- Customized YoloV5 pretrained model that was trained for face detection.
- Developed and examined a reliable binary image classification model for drowsiness detection using vision transformers.
- Applied real-time custom dataset testing with various scenarios to the framework.
- This research report is divided into the following sections in chronological order. A summary of the most recent methods for detecting drowsiness is given in Section II. The paper's methodology is outlined in Section II in the following order: picture augmentation comes first, followed by image classification using YoloV5.
- A detailed explanation of the model's experimental analysis and test on a customized dataset is given in Section III. A comparative study and major conclusions are provided in Section IV for comprehension.
- This study report is finally concluded in Section V.

II. **RELATED WORKS**

Since driving is becoming a common and essential activity for many people, it is necessary to make thorough efforts to comprehend, identify, and predict people's driving behaviours.

Many trials have been conducted to identify aberrant driver behaviour, such as drowsiness, in attempt to meet this demand. Modern techniques have been applied recently to enhance the functionality of sleepiness detecting systems. Zuojin Li et al and colleagues [16] looked at vehicle-based techniques.

Using sensors mounted on the vehicle, scientists were able to gather several forms of sleepiness data, such as yaw angles and steering wheel angles. After examining the characteristics derived from yaw and steering wheel angles, time series data is used to compute estimated entropy features.

The study used back-propagation neural networks as the classifier's input to determine the driver's level of drowsiness and achieved an accuracy of 87.21%. The motorist is classified by the system as awake, sleepy, and extremely sleepy.

Recently, behavioural approaches have been employed to address the issues raised by physiological and vehicle-based approaches. Because behavioural approaches concentrate on the driver's face expression rather than the behaviour of the car, they are more trustworthy than vehicle-based methods. Conversely, because of their complexity, physiological approaches are rarely commonly employed even if they yield incredibly accurate results.

Sherif Said et al. [17] advanced behavioural approaches by proposing a sleepiness detection system that uses the Viola Jones algorithm to identify face and eye areas. The motorist receives an alert from the system when it detects their level of tiredness. This method was evaluated in both indoor and outdoor settings, demonstrating 82% and 72.8% of the results, respectively.

An algorithm designed by Feng You [18] requires the driver to undergo offline training before it can be utilized online. The face areas are identified using Dlib's CNN, and the eye aspect ratio is computed using Dlib's 68 point facial landmarks. The sleepiness identification technique consists of two stages: online monitoring to detect the driver's state online and offline training using SVM classifier. Following a comparison examination, the suggested algorithm's accuracy is 94.8%; nevertheless, the primary drawback is that the SVM classifier needs to be trained by end users, meaning each driver must have their input. Convolutional neural networks (CNN) were introduced in [15] for prediction, and generative adversarial networks (GAN) for data augmentation.

Following a thorough investigation, the authors came to the conclusion that CNNs had improved the model's accuracy and that the use of GAN had generated fresh data samples (useful photos).

A method consisting of two sequential systems is proposed by R Tamanani et al. [19 The output system employs the Haar cascade technique for face detection and preprocessing of real time input video stream data, while the input system uses the CNN LeNet architecture for feature extraction and picture classification. Utilising stratified 5-fold cross validation on UTA-RLDD, the model yielded average scores of 91.8%, 92.8%, and 92% for accuracy, precision, recall, and F1-score. On training, validation, and testing data on a custom dataset, the system has shown accuracy of 98%, 84%, and 88%, respectively.

Novel designs YoloV5 [20], [21], which have advantages over current frameworks, were investigated for our suggested framework. A custom dataset is used to test the proposed framework and examine the performance in real-time.

III. METHODOLGY

A. Data Collection

The core of any good sleepiness detection system is data collection. Real-time video footage of drivers is being recorded in order to test and train the sleepiness detection algorithm in this project. The data collected covers a wide range of driving behaviors, including alert states and scenarios where the driver exhibits signs of sleepiness, such as yawning or closing their eyes. Precise annotation of these instances is essential to training and evaluating the model accurately. This design ensures dependability and compatibility, allowing for efficient data collecting for additional analysis. Data acquisition calls for carefully positioning the camera module inside the car or in a controlled environment in order to acquire the driver's facial features. The camera starts to record continuously, making sure that there is enough light to make facial features more visible. The camera begins to record continually, ensuring that there is sufficient light to enhance the visibility of face characteristics. During this process, a number of indicators of driver fatigue are noted, including yawning, eye closure, and changes in facial expression. To accurately recognize patterns suggestive of tiredness, the drowsiness detection system needs to understand these circumstances. All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

First, about 1100 related datasets were prepared. Using the picture labelling programme Make sense AI, two types of behaviours were tagged, and the dataset was labelled in YOLO format for later training operations. YOLOv5 finally completed behaviour detection following training. Every participant's random frames are gathered and binary tagged according to whether they are yawning or sleeping. The faces in the dataset exhibit notable changes in perception, scale, and attitude. Therefore, it is appropriate to use this efficiency in real-world applications. The YoloV5 is trained and annotated using this image collection.

B. Algorithm

Ultralytics, a computer vision and artificial intelligence research company, developed the object detection algorithm known as YOLOv5. In this mode, the YOLO (You Only Look Once) algorithm—which is renowned for its ability to detect objects quickly and precisely—has been improved. The Yolov5 algorithm builds upon the original Yolo algorithm's capabilities by introducing a number of new features and refinements. The Yolov5 is primarily based on a deep neural network design that is very good at identifying objects in pictures. Using a single neural network, the set of rules predicts item borders and class probabilities. Because YOLOv5 does not require several levels to detect objects, as do older algorithms, it is faster and more accurate than those methods. All things considered, YOLOv5 is an accurate, fast, and precise object recognition technique that may be used in a variety of computer vision applications. Its architecture and training techniques have been optimized, making it a popular option for object detection tasks. which have enhanced its overall effectiveness across several datasets. Fig.1 Architecture of Yolov5.

If Pi,j is not zero and the target is inside the bounding box, then it will be equal to one. A commonly used metric to describe the Intersection Over Union (IOU) between the true and estimated boxes is $IOUT.rue\ .\ Predicted$

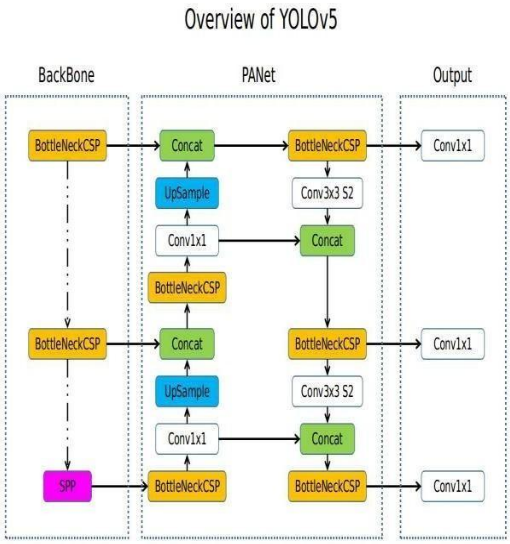


Fig. 1 Architecture of Yolov5

C. Face detection with YoloV5

To improve the model's performance, the wide-angle frames—which include most of the retrieved frames from both datasets—need to have their Region of benchmark dataset to illustrate performance and Interest (ROI) restored. The YoloV5 framework is used for face detection and automatic cropping of faces from wide-angle photos. The YoloV5 design combines the Darknet and the Cross Stage Partial Network (CSP). Annotating the ROI for a portion of the dataset's wide-angle frames yields the weights and configurations of the YoloV5 architecture. The annotated input image is used by the CSP and Darknet grid to get the features and target information. Face identification involves the creation of A × A grids from the input vector. In the event where the target's centre is a grid, then the target identification is the responsibility of that specific grid.

The location of the regression box of the face can be obtained by following (1):

$$C_{i,j} = P_{i,j} \times IOU_{Predicted\ True} \quad (1)$$

Where, C_i^j symbolises the confidence score of the jth bounding box of the ith grid. $P_{i,j}$ denotes whether there exists a target in the jth bounding box of the ith grid. The value $P_{i,j}$ indicates if a target is present in the jth bounding box.

IV. EXPERIMENTAL ANALYSIS AND RESULTS.

A. Dataset Preparation and Utilization.

This experiment makes use of a sizable dataset among When the IOU score is high, the expected box's location is predicted with more accuracy. The architecture of yoloV5 is shown in figure 1.
publicly accessible datasets for sleepiness estimate that we gathered from Kaggle and our own unique datasets. Different scenarios are included in the training dataset. The two most important scenarios are represented by the photos in the dataset: a mix of non-sleepiness-related activities (talking, laughing, staring at both sides) and symptoms of drowsiness (yawning, nodding). Every participant's random frames are gathered and binary labelled according to whether they are yawning or sleeping. The faces in the dataset exhibit notable changes in perception, scale, and attitude. Therefore, it is appropriate to use this benchmark dataset to illustrate performance and efficiency in real-world applications. The YoloV5 is trained and annotated using this image collection.

B. Computation Specifications of proposed system.

The hardware and software computational specs of our driver drowsiness detection framework are covered in this section. Python 3.9 was used to write the framework, which made use of Pytorch and OpenCV modules. A graphics processing unit with high performance is not used in the training of the YoloV5. Despite this, the framework maintains a high level of accuracy and efficiency. This is achieved by optimizing code and carefully selecting algorithms that balance speed with effectiveness. We also outline the minimum hardware specifications required for both training and testing the framework. Table I shows the specifications and minimal needs for testing and training the framework.

TABLE I

Specifications	System's Configuration
Operating system	Windows 11
CPU	Intel® i5 10th gen
RAM	15.8 Usable
GPU	Intel® UHD Graphics
Frameworks	Pytorch, OpenCV

C. Evaluating YoloV5 for face detection

The fundamental architecture of the YoloV5 is implemented with all the settings needed for precise face detection and ROI extraction from the input image. The model was trained using 100 epochs with specific settings. The model's lowest confidence score, as assessed by many tests, is 0.75. Fig.3 shows the accuracy, recall, and mean average precision metric (mAP) at 0.5 threshold curves. Fig.4 displays the trained architecture's detection effect on dataset samples. Following a battery of experiments, the trained YoloV5 achieved a speed of 51.9 pictures per second. The ROI is cut from the wide-angle frames of the training and validation frames once the bounding box positions are established.

96.2% and 97.4%, respectively. These results are obtained with a particular set of hyper-parameters, as indicated in Table II.

Fig. 2 Results after detection

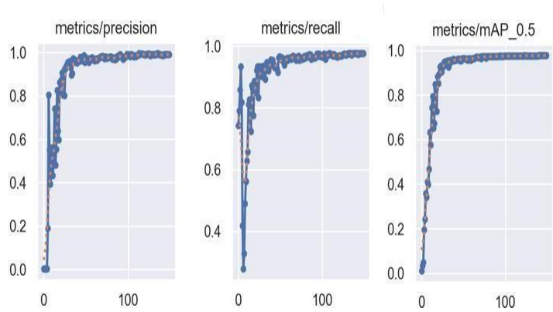
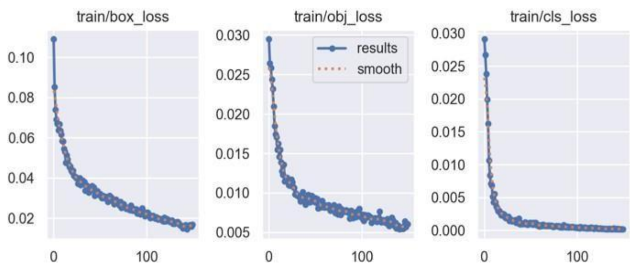
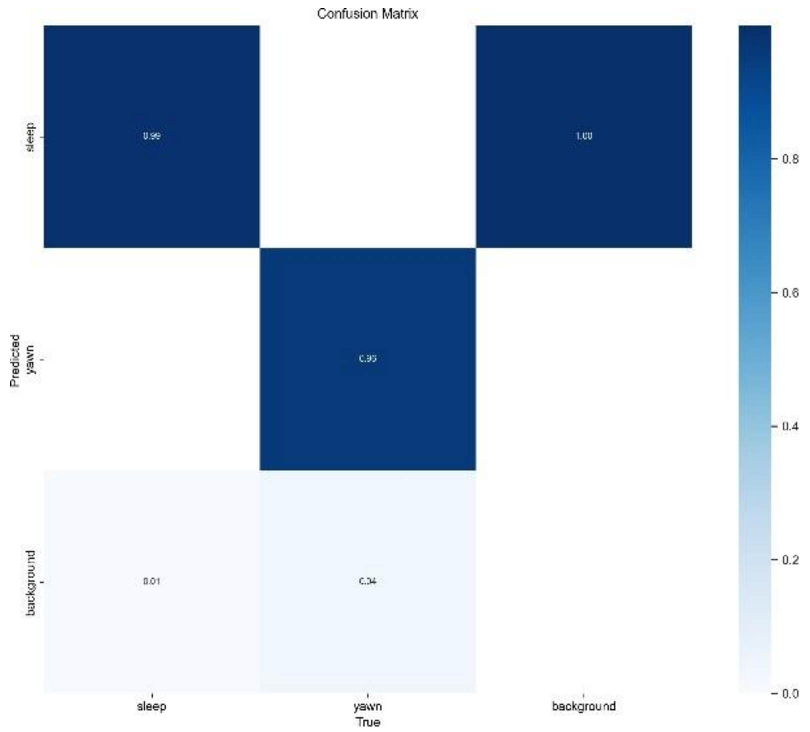


Fig. 3 Visualization of YoloV5 Performance for 100 epochs.

Fig. 4 Visualization of validation loss.





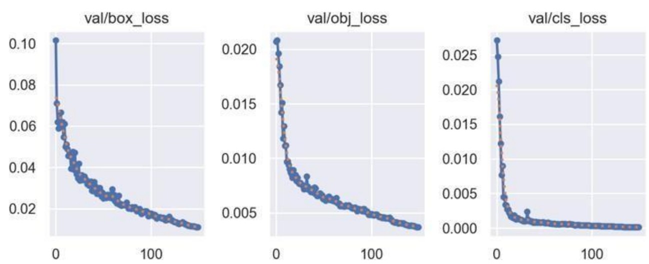


Fig. 5 Visualization of Training loss.

Fig. 6 Confusion matrix and learning curves of training and validation accuracy and loss for 100 epochs.

TABLE III

TUNED HYPERPARAMETERS OF YOLOV5

Hyper-parameter	**Attribute**
No. of classes	2
Input shape of image	(256, 256, 3)
Resized image Size	(392, 392)
Patch Size	30
Batch Size	264
Number of Epochs	100
Learning Rate	0.001
Weight Decay	0.0001
Number of Heads	4

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Where, F_P is the quantity of non-drowsy images that were incorrectly identified as the state of drowsiness; T_n is the number of non-drowsy images that are accurately identified as being in the non-drowsiness state. Where T_P is the number of photographs that were correctly detected as being in the drowsiness state, and F_n is the number of photos that were mistakenly identified as being in the non-drowsiness state. The calculated precision, recall, and F1-score values are shown in Table No. III.

TABLE IIIII

CALCULATED EVALUATION METRICS

Researched By	Facial Detector	Overall Accuracy	Testing in real time
Bakheet et al. [12]	Haar Cascades	85.62 %	×
Shreyans M et al. [27]	Dlib	75.67 %	×
R Tamanani et al. [19]	Haar Cascades	91.8 %	C
Anh-Cang et al. [28]	SSD Network	97%	×
Proposed Framework	**YoloV5**	**97.4 %**	C

Precision, recall, and F1 scores, among other accuracy metrics, were estimated and displayed. The following.

D. Evaluating trained model.

Effective techniques are added to the training set after face detection. The YoloV5 model is trained in this experiment to classify sleepiness states as either alert or drowsy. If the user is sleepy or yawns for longer than three seconds, an alert will sound. The accuracy and loss learning curves for the models' training and validation are displayed. The validation and training curves in these learning graphs both maintain a point of stability with little gap, indicating a well-fitting learning algorithm. To optimize the model's performance, three missions were used at the same time for training.

1) Output computation; 2) error debugging; and 3) hyper- parameter adjustment. The maximum training and validation accuracy after several iterations of hyper-parameter tuning are equations (2), (3), and (4) are used to evaluate the precision, sensitivity (recall), and F1-score, respectively.

variety of scenarios the model is exposed to during training, the way data is annotated can have a major effect on the model's capacity to generalize.

More robust models can be produced as a result of proper data annotation, which guarantees a wide representation of cases and improves training results. When creating machine learning models, these insights should be kept in mind to guarantee high accuracy across various datasets and testing settings.

TABLE IVV

COMPARISION OF OUR PROPOSED FRAMEWORK WITH THE EXISTING MODELS

State	Precision	Recall	F1 Score	Image Support
Yawn	0.97	0.98	0.97	626
Sleep	0.96	0.98	0.98	450

Furthermore, we observed notable variations in accuracy when adjusting the train and validation split ratios, suggesting that this variable is crucial in influencing the model's testing accuracy. With training accuracy reaching 97.7%, validation accuracy reaching 98.3%, and testing accuracy reaching 95.5%, we discovered that an 80-20 split yielded the best results. This conclusion emphasizes how crucial it is to strike the ideal balance when splitting the data, since it can have a direct impact on the model's functionality. Furthermore, it's critical to consider the quality of the datasets utilized for validation and training. Because it influences the number and time sleepiness detection that has never been used before. Several experiments are conducted using the custom real-time dataset.

E. Comparison with existing models.

To achieve the best results, it's crucial to combine the right architectures for face detection and categorization. Numerous frameworks are offered with various computer vision and machine learning architectures in order to attain optimal performance. This subsection's main goal is to analyze the most egregious research endeavors that aim to use machine learning architectures to detect human sleepiness. Table VIII presents a thorough comparison analysis of human sleepiness classification methodologies. In recent sleepiness detection research endeavors, numerous face identification algorithms have been created, ranging from CNNs to haar cascades. Numerous image classification algorithms, including CNNs and Bayesian classifiers, had also been applied in these endeavors.

V. AUTHOR CONTRIBUTIONS: FINDINGS AND COMPARATIVE ANALYSIS.

This section highlights the important discoveries and compares our suggested model to previous studies that have already been done in this area. This section compares our suggested framework to two different criteria: the technique and the dataset. It consists of both public and own custom datasets and YoloV5 algorithm is utilized.

Dataset: Several characteristics in the dataset to be analyzed play a major role in the ability of driver sleepiness detection systems. There is an enormous amount of data involved, which suggests that assessing the system in real-time is more challenging [26] [29]. This is also because the type of image data that is available in public databases is unclear. We collected a customized dataset based on real-time requirements in order to overcome the aforementioned problems, and this produced the best testing outcomes. Additionally, this dataset is easy to examine, and there's a strong chance it will be enlarged as more researchers on sleepiness detection systems around the world may add a lot.

YoloV5: Conventional neural networks (CNN) [13], Generative Adversarial Networks (GAN) [15], and conventional computer vision algorithms [11] are the mainstays of contemporary research on sleepiness detection. We have put up a brand-new YoloV5 framework for real- time sleepiness detection that has never been used before. Several experiments are conducted using the custom real-time dataset.

VI. CONCLUSION.

We have utilized YoloV5 for effective driver drowsy state estimation in this study. This framework consists mostly of two successive components: the conclusive component does binary image classification, and the initial component uses YoloV5 pre-trained face detection CNN architecture to execute automatic face cropping. Following multiple comprehensive analyses, the YoloV5 attained a mAP score of almost 95%. The model's assessment showed that the framework had achieved high average precision, sensitivity, and F1-score values—0.97, 0.98, and 0.97, as appropriate. Using a custom dataset also shows that the model achieved 95.5% accuracy throughout testing.

The following is a summary of the suggested architecture's limitations. First off, training the suggested model requires a large quantity of data with labelled sequence conditions, even if it achieved good detection performance. Future research may consider a number of arguments. In order to lower costs and boost computational efficiency without degrading, we will first optimize the network configuration in the suggested design for usage in microcomputer systems. In order to improve the model's performance, we will secondly employ generative adversarial networks to expand the amount of training data.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to our guide, Prof. Asutosh Pradhan, for his valuable guidance, constant support, and encouragement throughout the project. We are also thankful to RD Engineering College, Ghaziabad, for providing us with the necessary facilities and resources to complete our work. Our heartfelt thanks to our peers and contributors for their feedback, and to all researchers whose work inspired and shaped our study.

REFERENCES

[1] "Nhtsa report(accessed on 4 august 2021)." [Online]. Available: https://www.nhtsa.gov

[2] J. Yu, S. Park, S. Lee, and M. Jeon, "Driver drowsiness detection using condition-adaptive representation learning framework," *IEEE Transac- tions on Intelligent Transportation Systems*, vol. 20, no. 11, pp. 4206– 4218, 2019.

[3] Y. S. S. E. Viswapriya, Singamsetti Balabalaji, "A machine-learning approach for driver-drowsiness detection based on eye-state," *INTER- NATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOL-OGY (IJERT)*, vol. 10, April 2021.

[4] M. M. Hasan, C. N. Watling, and G. S. Larue, "Physiological signal- based drowsiness detection using machine learning: Singular and hybridsignal approaches," *Journal of Safety Research*, 2021.

[5] D. K. D. K.Mirunalini, "Drowsiness detection using deep neural network," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 9, pp. 317–326, 2021.

[6] M. Gromer, D. Salb, T. Walzer, N. M. Madrid, and R. Seepold, "Ecg sensor for detection of driver's drowsiness," *Procedia Computer Science*, vol. 159, pp. 1938–1946, 2019, knowledge-Based and Intelligent Infor- mation & Engineering Systems: Proceedings of the 23rd International Conference KES2019.

[7] M. Zhu, H. Li, J. Chen, M. Kamezaki, Z. Zhang, Z. Hua, and S. Sugano, "Eeg-based system using deep learning and attention mechanism for driver drowsiness detection," in *2021 IEEE Intelligent Vehicles Sympo- sium Workshops (IV Workshops)*, 2021, pp. 280–286.

[8] M. Song, L. Li, J. Guo, T. Liu, S. Li, Y. Wang, Qurat ul ain, and J. Wang, "A new method for muscular visual fatigue detection using electrooculogram," *Biomedical Signal Processing and Control*, vol. 58, p. 101865, 2020.

[9] L. R. Femilia Hardina Caryn, "Driver drowsiness detection based on drivers' physical behaviours: A systematic literature review," *Computer Engineering and Applications*, vol. 10, no. 3, 2021.

[10] H. U. R. Siddiqui, A. A. Saleem, R. Brown, B. Bademci, E. Lee, F. Rustam, and S. Dudley, "Non-invasive driver drowsiness detection system," *Sensors*, vol. 21, no. 14, 2021.

[11] L. Thulasimani, P. P, and P. S P, "Real time driver drowsiness detectionusing opencv and facial landmarks," *Int. J. of Aquatic Science*, vol. 12, no. 2, pp. 4297–4314, 2021.

[12] S. Bakheet and A. Al-Hamadi, "A framework for instantaneous driver drowsiness detection based on improved hog features and na¨ıve bayesianclassification," *Brain Sciences*, vol. 11, p. 240, 02 2021.

[13] A. A. Jordan, A. Pegatoquet, A. Castagnetti, J. Raybaut, and P. Le Coz, "Deep learning for eye blink detection implemented at the edge," *IEEE Embedded Systems Letters*, vol. 13, no. 3, pp. 130– 133, 2021.

[14] T. Vesselenyi, S. Moca, A. Rus, T. Mitran, and B. Taˇtaru, "Driver drowsiness detection using ANN image processing," *IOP Conference Series: Materials Science and Engineering*, vol. 252, p. 012097, oct 2017.

[15] M. Ngxande, J.-R. Tapamo, and M. Burke, "Bias remediation in driver drowsiness detection systems using generative adversarial networks," *IEEE Access*, vol. 8, pp. 55 592–55 601, 2020.

[16] Z. Li, C. Liukui, J. Peng, and Y. Wu, "Automatic detection of driver fatigue using driving operation information for transportation safety," *Sensors*, vol. 17, p. 1212, 05 2017.

[17] S. Said, S. Alkork, T. Beyrouthy, M. Hassan, O. E. Abdellatif, and M. F. Abdraboo, "Real Time Eye Tracking and Detection- A Driving Assistance System," *Advances in Science, Technology and Engineering Systems Journal*, vol. 3, no. 6, pp. 446–454, 2018.

[18] F. You, X. Li, Y. Gong, H. Wang, and H. Li, "A real-time driving drowsiness detection algorithm with individual differences considera- tion," *IEEE Access*, vol. 7, pp. 179 396–179 408, 2019.

[19] R. Tamanani, R. Muresan, and A. Al-Dweik, "Estimation of driver vigilance status using real-time facial expression and deep learning," *IEEE Sensors Letters*, vol. 5, no. 5, pp. 1–4, 2021.

[20] Z. X. Xu Q Zhu Z, Ge H Zhang Z, "Effective face detector based on yolov5 and superresolution reconstruction." *Comput Math Methods Med.2021*, Published 2021 Nov 16.

[21] D. Qi, W. Tan, Q. Yao, and J. Liu, "Yolo5face: Why reinventing a face detector," 2022.

[22] A. Dosovitskiy, L.Beyer, A.Kolesnikov, D.Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Trans- formers for image recognition at scale," *CoRR*, vol. abs/2010.11929, 2020.

[23] H. Dong, L. Zhang, and B. Zou, "Exploring vision transformers for polarimetric sar image classification," *IEEE Transactions on Geoscienceand Remote Sensing*, pp. 1–1, 2021.

[24] P. Deng, K. Xu, and H. Huang, "When cnns meet vision transformer: A joint framework for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[25] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision transformers for remote sensing image classification," *Remote Sensing*, vol. 13, no. 3, 2021.

[26] R. Ghoddoosian, M. Galib, and V. Athitsos, “A realistic dataset and baseline temporal model for early drowsiness detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.

[27] S. Mittal, S. Gupta, Sagar, A. Shamma, I. Sahni, and N. Thakur, “Driver drowsiness detection using machine learning and image processing,” in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021, pp. 1–8.

[28] A.-C. Phan, N.-H.-Q. Nguyen, T.-N. Trieu, and T.-C. Phan, “An efficientapproach for detecting driver drowsiness based on deep learning,” *Applied Sciences*, vol. 11, no. 18, 2021.

[29] S.-H. L. Ching-Hua Weng, Ying-Hsiu Lai, “Driver drowsiness detectionvia a hierarchical temporal deep belief network,” *In Asian Conferenceon Computer Vision , Taipei, Taiwan*, Nov. 2016

.