iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50345

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Driver Drowsiness Detection System using Deep Learning

Devarakonda Sruthi[1], Avanaganti Amulya Reddy[2], G. Sai Siddaharth Reddy[3], Mrs. Shilpa Shesham[4]
[1, 2, 3]UG scholar, [4]Assistant Professor, Department of AI, Anurag Group Of Institutions, Hyderabad.

Abstract: ***These days, an ever-increasing number of professions require long time focus. Drivers should watch out for the street, so they can respond to abrupt occasions right away. Due to driving for a long time or intoxication, drivers might feel sleepy, which is the biggest distraction for them while driving. This distraction might cost the death of the driver and other passengers in the vehicle, and at the same time, it also causes the death of people in the other vehicles and pedestrians too. To prevent such accidents, we propose a system that helps to alert the driver if he/she feels drowsy. To accomplish this, we implement the solution using a computer-vision-based machine learning model. The driver's face is detected by a face recognition algorithm continuously using a camera, and the face of the driver is captured. The face of the driver is given as input to a classification algorithm which is trained with a data set of images of drowsy and non-drowsy faces. The algorithm uses landmark detection to classify the face as drowsy or not drowsy. If the driver's face is drowsy, a voice alert is generated by the system. This alert can make the driver aware that he/she is feeling drowsy, and the necessary actions can then be taken by the driver. This system can be used in any vehicle on the road to ensure the safety of the people who are traveling and prevent accidents that are caused due to the drowsiness of the driver.***

Keywords: Computer Vision, Deep Learning, Convolutional Neural Network, Eye Aspect Ratio, Mouth Aspect Ratio.

I. INTRODUCTION

Accidents due to driver drowsiness are a significant problem worldwide. When drivers are tired or sleepy, their ability to react and make quick decisions is impaired, and they may even fall asleep at the wheel, resulting in accidents. According to the World Health Organization, driver fatigue is estimated to cause up to 20% of road accidents globally. Statistics from various countries highlight the seriousness of the problem. There are typically three primary techniques used to identify drowsiness:

1) *Behavioural Parameter-Based Techniques:* Behavioural parameters are non-invasive measures for drowsiness detection. These techniques measure driver's fatigue through behavioural parameters of the driver, such as eye closure ratio, eye blinking, head position, facial expressions, and yawning. The Percentage of Eye Closures (PERCLOS) are one of the most commonly used metrics in detecting drowsiness based on eye state observation. PERCLOS is the ratio of eye closure over a period, and then on the result of PERCLOS, eyes are referred to as open or closed. Yawning-based detection systems analyse the variations in the geometric shape of the mouth of a drowsy driver, such as the broader opening of the mouth, lip position, etc. Behavioural-based techniques use cameras and computer vision techniques to extract behavioural features.
2) *Vehicular Parameters-Based Techniques:* Vehicular parameter-based methods try to detect driver fatigue based on vehicular features such as frequent lane-changing patterns, vehicle speed variability, steering wheel angle, steering wheel grip force, etc. These measures require sensors on vehicle parts like the steering wheel, accelerator, brake pedal, etc. The signals generated by these sensors are used to analyse drivers' drowsiness. The main goal of these techniques is to observe driving patterns and detect a decline in driving performance due to fatigue and tiredness.
3) *Physiological Parameters-Based Techniques:* The Physiological parameters-based methods detect drowsiness based on drivers' physical conditions such as heart rate, pulse rate, breathing rate, respiratory rate, body temperature, etc. Fatigue or drowsiness changes physiological parameters such as decreased blood pressure, heart rate, body temperature, etc. Physiological parameters-based drowsiness detection systems detect these changes and alert the driver when he is in the state, near to sleep. The advantage of this approach is that it alerts the driver to rest before the physical symptoms of drowsiness appear.

A driver drowsiness detection system is a technology that uses various sensors, algorithms, and artificial intelligence to monitor the driver's behaviour and detect signs of drowsiness or fatigue. The system can issue an alert to the driver through an audio warning or any other alert to prevent accidents before they occur. One of the most popular and effective driver drowsiness detection approaches is computer vision and deep learning techniques.

©IJRASET: All Rights are Reserved

Computer vision involves using cameras and image processing algorithms to capture and analyse the driver's facial features, such as eye movements and facial expressions, to detect signs of drowsiness.

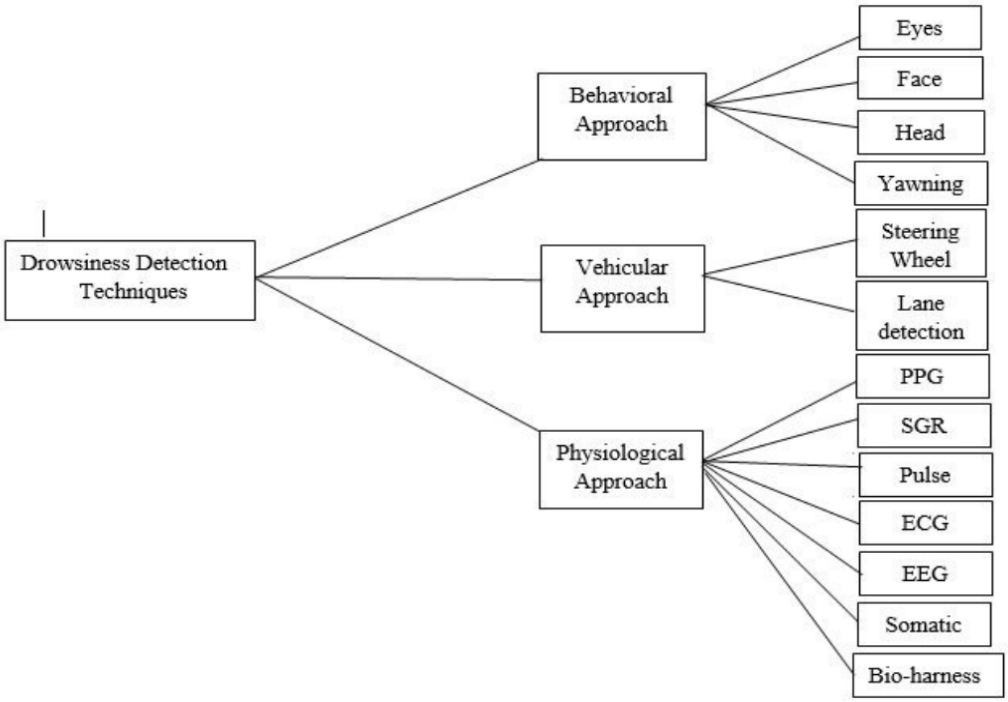


Figure: Drowsiness detection techniques.

To identify patterns related to drowsiness, deep learning algorithms like CNN may be trained on a big dataset of images. The combination of CV and DL has led to the development of advanced driver drowsiness detection systems that can detect the driver's level of fatigue in real-time. These systems can be integrated into vehicles or installed as an aftermarket product, making them accessible to many drivers. Driver drowsiness detection systems can prevent accidents and save lives by alerting drivers before they become too fatigued to operate a vehicle safely. These systems can be especially beneficial for commercial drivers, such as truck drivers, at higher risk of drowsy driving due to long working hours and inadequate rest breaks. In reality, driver drowsiness detection systems are an important technological development in road safety, and their use will likely increase. As the technology continues to evolve, it may become more accurate and accessible, preventing accidents due to driver fatigue.

II. EXISTING SYSTEMS

A. *Principal Component Analysis (PCA)*

PCA is a popular dimensionality reduction technique in machine learning and data analysis. It is commonly used for feature extraction and data visualization, transforming high-dimensional data into a lower-dimensional space while preserving important information. However, PCA also has certain disadvantages in the context of driver drowsiness detection systems, which are used to alert drivers when they show signs of falling asleep while driving.

Disadvantages of using Principal Component Analysis (PCA) for driver drowsiness detection systems:

1) May not effectively capture subtle changes in facial expressions and eye movements associated with drowsiness.
2) May not be robust to inter-subject variability, leading to reduced performance in real-world scenarios.
3) Requires a sufficient amount of labeled data for training, which may be challenging to obtain for drowsiness episodes during driving.
4) Assumes linear relationships between variables, but the relationship between facial expressions, eye movements, and driver drowsiness may be nonlinear.
5) Limited ability to capture complex nonlinearities and individual differences in drowsiness patterns.
6) Linear nature of PCA may not fully capture the complexity of drowsiness patterns, leading to reduced accuracy and reliability.
7) Potential risk and ethical concerns in collecting labeled data of drivers experiencing genuine drowsiness episodes while driving.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Support Vector Machines (SVM)

SVM is a popular machine learning algorithm for classification and regression tasks. It works by finding a hyperplane that best separates data points of different classes or predicts the target variable for regression while maximizing the margin between the classes. SVM has been widely used in various applications, including image recognition, speech recognition, bioinformatics, and finance.

Disadvantages of using Support Vector Machines (SVM) for driver drowsiness detection systems:

1) SVM is a binary classification technique that may not be suitable for detecting subtle levels of drowsiness that may exist on a continuous spectrum.
2) SVM may not effectively capture driver drowsiness's temporal dynamics and time-varying patterns, as it primarily operates on static feature vectors.
3) SVM's performance may degrade in real-world scenarios due to environmental factors such as lighting conditions, driver pose variations, and motion artifacts that can affect the accuracy of feature extraction.
4) SVM may be unable to effectively handle inter-subject variability, as different individuals may exhibit different patterns of facial expressions and eye movements when drowsy.
5) SVM may not be well-suited for handling imbalanced datasets, which may be common in driver drowsiness detection scenarios where drowsy episodes are relatively rare compared to non-drowsy episodes.
6) SVM's ability to handle noisy data may be limited, as it is sensitive to outliers and misclassifications, which can occur in real-world driving scenarios.
7) SVM may not fully capture the nonlinear relationships between facial expressions, eye movements, and driver drowsiness, as it relies on the linear separation of data points in feature space.
8) Deployment of SVM-based drowsiness detection systems in real-world vehicles may pose challenges regarding computational complexity and real-time processing requirements.
9) Ethical concerns and potential risks in collecting labelled data of drivers experiencing genuine drowsiness episodes while driving for training an SVM model.

III. PROPOSED METHODOLOGY

This section will discuss the proposed methodology and techniques. The dataset for this work is taken from the open-source website, and the dataset is called the yawn_eye_dataset, available on Kaggle. The yawn_eye_dataset contains around 3000 RGB images. This dataset comes with two different folders, train and test, which are divided into four folders, i.e., open, closed, yawn, and no_yawn. In the proposed methodology, there are four stages.

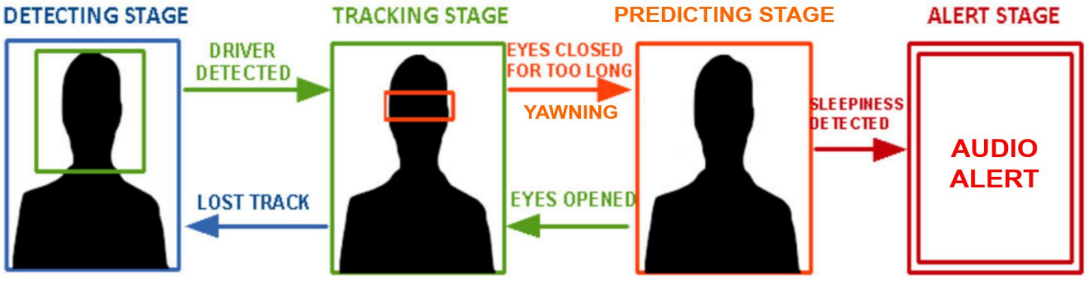


Figure: Methodology

A. Detecting Stage

Driver drowsiness detection is an important application of computer vision and deep learning techniques. One of this project's initial stages is detecting the driver's face. This is typically done using face detection algorithms, which can detect the location and size of the face in an image or video frame. Haar cascades are a machine learning-based approach to object detection, which uses Haar-like features and a cascading classifier to detect objects in images or videos. Haar-like features are simple rectangular features that are used to represent local image properties. A cascading classifier is a series of classifiers trained to detect increasingly complex features of an object, with each Stage of the cascade reducing the number of false positives. Driver's face in real-time is detected by using OpenCV. The OpenCV's inbuilt features, i.e., Haar feature-based cascade classifiers. The following cascade is used to classify the input and to detect the face of the driver.

1) haarcascade_frontalface_default.xml

The face detection stage is thus a crucial step in the overall driver drowsiness detection pipeline, as it provides the foundation for subsequent analysis of the driver's behaviour.

B. Tracking Stage

The tracking stage involves selecting the relevant area, i.e., the Region of Interest (ROI) of the image or video frame where the driver's eyes and mouth are located. This is typically done after the face detection stage, which identifies the location of the driver's face. The ROI is important because it provides the specific area of the image or video frame that needs to be analysed for signs of drowsiness, such as eye closure or prolonged periods of eye fixation or yawning. Creating an accurate ROI requires careful consideration of factors such as camera position, lighting conditions, and the driver's posture. It is also important to account for variations in the driver's position and orientation over time, as well as the presence of other objects in the image that may interfere with face detection. The ROI is typically created by using face detection algorithms, such as Haar cascades and face landmarks. Such as:

1) shape_predictor_68_face_landmarks.dat
2) haarcascade_frontalface_default.xml

C. Predicting Stage

In this Stage, the ROI, i.e., eyes and mouth, are fed to the Classifier. The Classifier will categorize whether the eyes and mouth are open or closed. In the Proposed methodology, a well-trained CNN acts as the Classifier. Convolutional Neural Networks (CNN) are chosen as the deep learning methodology for the development of the Classifier. Four convolutional layers are added to this model, along with the Max pooling layer, Batch Normalization, and dropout layer. Batch Normalization is used to accelerate and make the network stale during the training of deep neural networks. Batch normalization offers some regularization effect, reducing generalization error. The preferred approach to minimize neural network overfitting is to employ dropout layers. Higher-level features are extracted from raw image pixel data by CNNs using various filters, which the model then uses to classify the data.

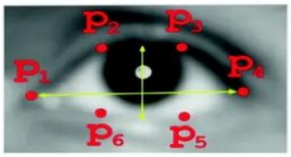
CNN includes three segments: Convolutional layers, which employ a particularized number of convolution filters to the image. The layer performs a set of mathematical processes for each sub-region to produce a single mark in the output feature map. Convolutional layers then typically implement a ReLU activation function to the output. A regularly used pooling algorithm is max pooling, which extracts sub-regions of the feature map, keeps their greatest value, and discards all other values. Dense or fully connected layers perform classification on these feature maps. In a dense layer, every node in the layer is joined to every other node in the previous layer. When compiling the model, categorical_crossentropy is chosen as the loss function and Adam optimizer.

```
Model: "sequential_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_8 (Conv2D)            (None, 222, 222, 16)      448
_________________________________________________________________
max_pooling2d_8 (MaxPooling2 (None, 111, 111, 16)      0
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 111, 111, 32)      4640
_________________________________________________________________
max_pooling2d_9 (MaxPooling2 (None, 55, 55, 32)        0
_________________________________________________________________
batch_normalization_8 (Batch (None, 55, 55, 32)        128
_________________________________________________________________
dropout_8 (Dropout)          (None, 55, 55, 32)        0
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 53, 53, 64)        18496
_________________________________________________________________
max_pooling2d_10 (MaxPooling (None, 26, 26, 64)        0
_________________________________________________________________
batch_normalization_9 (Batch (None, 26, 26, 64)        256
_________________________________________________________________
dropout_9 (Dropout)          (None, 26, 26, 64)        0
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 24, 24, 128)       73856
_________________________________________________________________
max_pooling2d_11 (MaxPooling (None, 12, 12, 128)       0
_________________________________________________________________
batch_normalization_10 (Batc (None, 12, 12, 128)       512
_________________________________________________________________
dropout_10 (Dropout)         (None, 12, 12, 128)       0
_________________________________________________________________
flatten_2 (Flatten)          (None, 18432)             0
_________________________________________________________________
dense_4 (Dense)              (None, 128)               2359424
_________________________________________________________________
batch_normalization_11 (Batc (None, 128)               512
_________________________________________________________________
dropout_11 (Dropout)         (None, 128)               0
_________________________________________________________________
dense_5 (Dense)              (None, 4)                 516
=================================================================
Total params: 2,458,788
Trainable params: 2,458,084
Non-trainable params: 704
```

Figure: CNN Model Summary

Eye Aspect Ratio (EAR)

EAR is a widely used metric for measuring eye-opening and is commonly used in facial expression analysis, eye tracking, and driver drowsiness detection systems. EAR is calculated by measuring the ratio of the distance between the vertical landmarks of the eye (the upper and lower eyelids) to the distance between the horizontal landmarks of the eye. The EAR calculation is based on the fact that when a person's eyes are open, the distance between the upper and lower eyelids will be greater than the distance between the inner and outer corners of the eye. Conversely, when the eyes are closed, the distance between the eyelids will decrease, leading to a decrease in the EAR value.



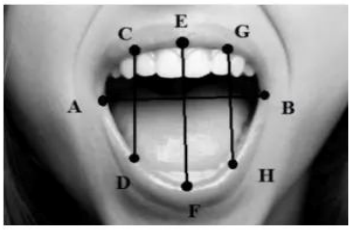
$$\mathrm{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Figure: Eye Aspect Ratio Formula

Where p1, p2, p3, p4, p5, and p6 are the six landmark points corresponding to the eye. Specifically, p1 and p4 are the landmarks at the inner and outer regions of the eye, respectively, and p2, p3, p5, and p6 are the landmarks at the upper and lower eyelids. If the EAR value falls below a certain threshold, it may be an indication that the eyes are partially or completely closed, which could be a sign of drowsiness or fatigue. By continuously monitoring the EAR value, these systems can alert drivers when they feel drowsy and helps to prevent accidents caused by driver fatigue. The threshold value of 0.3 for EAR is often used in driver drowsiness detection systems. When a person's eyes are fully open, the EAR value is typically around 0.3. As the eyes close, the EAR value decreases, and values below 0.3 indicate that the eyes are partially or fully closed. A threshold value of 0.3 is also considered to be a conservative value, meaning that it errs on the side of caution and is less likely to miss instances of drowsiness or fatigue. Using a higher threshold value may result in missing instances of drowsiness, while using a lower threshold value may result in false alarms or unnecessary warnings.

D. Mouth Aspect Ratio (MAR)

The MAR is a measure of the mouth opening and is commonly used in facial expression analysis and emotion detection. MAR is calculated by measuring the ratio of the distance between the vertical landmarks of the mouth (the upper and lower lips) to the distance between the horizontal landmarks of the mouth (the corners of the mouth). MAR can be used to detect various facial expressions, such as smiles or frowns, as well as to detect emotions, such as happiness or sadness. The calculation of MAR is based on the assumption that when a person's mouth is open, the distance between the upper and lower lips will be greater than the distance between the corners of the mouth. Conversely, when the mouth is closed, the distance between the lips will decrease, leading to a decrease in the MAR value.



$$\mathrm{MAR} = \frac{|EF|}{|AB|}$$

Figure: Mouth Aspect Ratio Formula

Where E and F are the vertical landmarks at the upper and lower lips, respectively, and A and B are the horizontal landmarks at the corners of the mouth. If the MAR value falls below a certain threshold, it may be an indication that the person's mouth is closed or partially closed, which could be a sign of sadness, stress, or lack of alertness. MAR can also be used in conjunction with EAR (Eye Aspect Ratio) to detect drowsiness or fatigue. If both the EAR and MAR values fall below their respective thresholds, it may be an indication that the person is experiencing drowsiness or fatigue.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

E. Alert Stage

After the model is trained with the given dataset, we can use this model to predict the class of the images which are captured from the camera. We use OpenCV to capture the images from the camera. We continuously capture image frames from the camera. The same pre-processing steps which are applied on the dataset are applied on each frame captured, i.e., detecting the face from the image frame, extracting the Region of Interest, and then resizing the Region of Interest to a fixed size. Then we convert the images into array format to give as input to the model. Then, we can give a set of images to the trained CNN Classification model to predict the labels for the images. Once the labels are predicted then their EAR and MAR values will be calculated. The alert audio Stage is activated when the EAR or MAR values fall below a certain threshold, indicating that the driver is feeling drowsy. Once this threshold is crossed, the system triggers an audio alert, which can be in the form of a loud beep, a voice command, or a sound signal. The audio alert is designed to grab the driver's attention and prompt them to take corrective action, such as opening their eyes wider, adjusting their posture, or taking a break. In conclusion, the final alert audio Stage is a critical component of the Driver Drowsiness Detection system, designed to ensure that the driver remains alert and attentive throughout their journey.

IV. RESULTS

A. CNN Results

After an extensive training process on a large dataset, the CNN model has achieved impressive results in terms of accuracy. The CNN model's superior performance in both training and testing phases validates its effectiveness as a powerful classifier, capable of accurately categorizing data into appropriate. The model's consistent and impressive results highlight its reliability and suitability for real-world scenarios, making it a promising choice for diverse machine learning and artificial intelligence applications.

B. Training Results

In the training phase after training the proposed model on the training dataset, these are the results which we have obtained. The highest training accuracy is observed at 80 epochs.

Table: Training Results

Epochs	Training Accuracy	Training Loss
10	97.53	6.26
20	97.97	5.39
30	98.99	3.05
40	98.62	3.38
50	99.07	2.65
60	98.09	5.47
70	99.51	1.12
80	99.84	0.60
90	99.59	0.13
100	99.39	1.94

©IJRASET: All Rights are Reserved

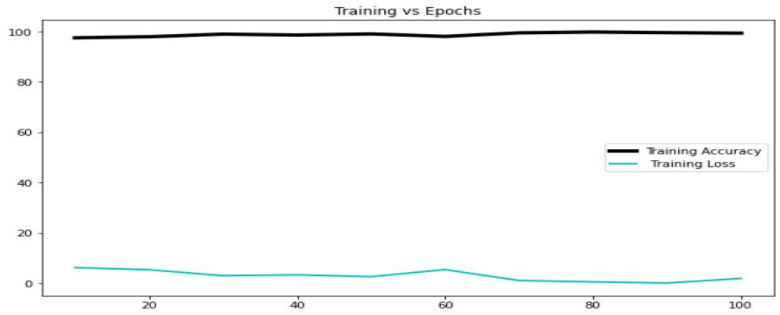


Figure: Visualization of Training Results

C. Testing Results

These are the accuracies which we have obtained and the highest Testing accuracy is observed at 80 epochs.

Table: Testing Results

Epochs	Training Accuracy	Testing Accuracy
10	97.53	95.84
20	97.97	97.69
30	98.99	95.84
40	98.62	97.69
50	99.07	97.92
60	98.09	97.69
70	99.51	97.23
80	99.84	99.31
90	99.59	97.69
100	99.39	96.30

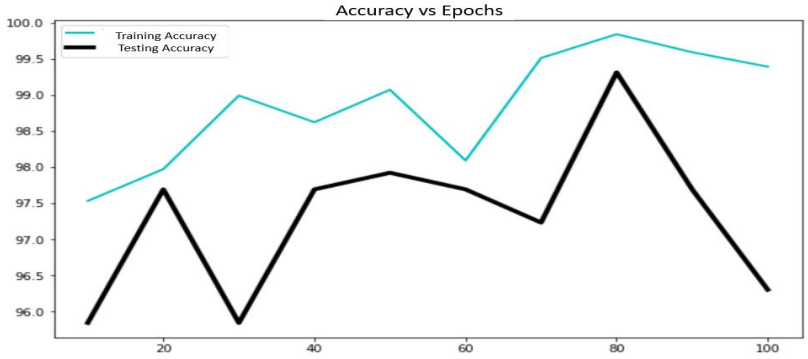


Figure: Visualization of Testing Results

D. Result Analysis

After careful evaluation, it can be concluded that the proposed Convolutional Neural Network (CNN) methodology outperforms previous works. The CNN model demonstrates higher accuracy and efficiency in classification tasks, showcasing its superiority in handling complex data.

Table: Comparison Table

Reference Paper No.	Accuracy
1	96
2	97
3	96
Proposed Methodology	99

Results of frames captured from camera:

1) Frames Classified As Drowsy

Here the input will be the continuous stream of video. EAR and MAR values are continuously tracked. A text message will be prompted on the screen and an audio alert is activated when the EAR or MAR values fall below a certain threshold, indicating that the driver is feeling drowsy. Once this threshold is crossed, the system triggers an audio alert, which can be in the form of a loud beep, a voice command or a sound signal.

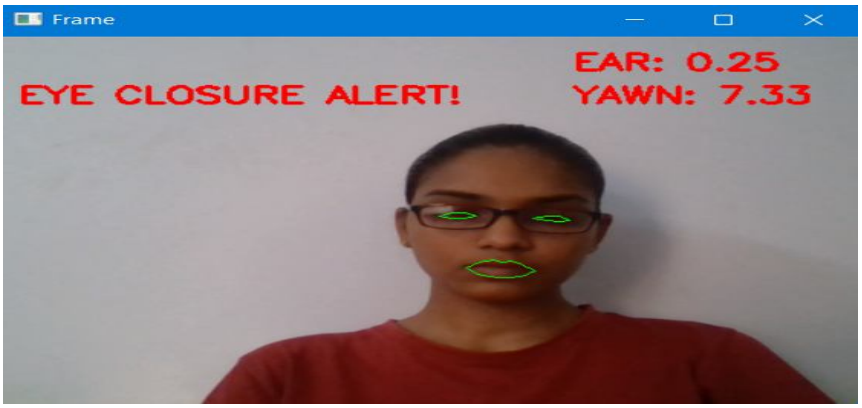


Figure: Drowsy Output 1

In the above output drowsiness is detected because the person had closed the eyes for too long, and the EAR value falls below the eye threshold value. Because of which the system has classified that the person is feeling drowsy.

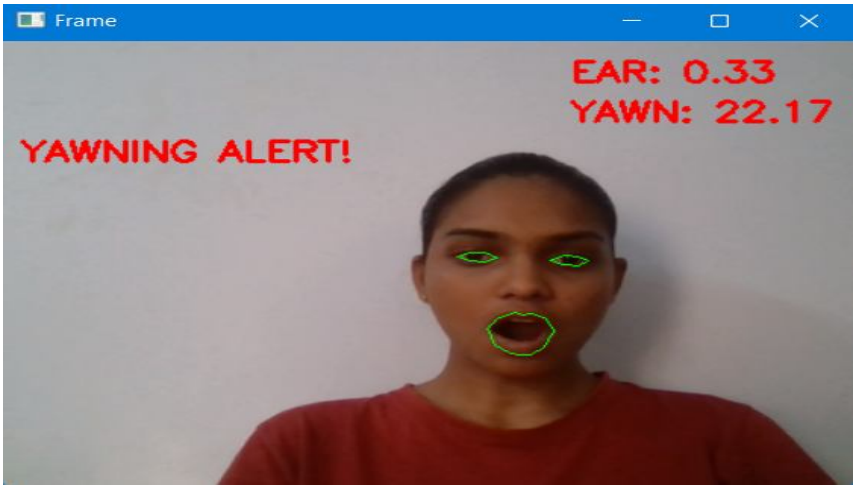


Figure: Drowsy Output 2

In the above output drowsiness is detected because the person has Yawned, and the MAR value falls below the mouth threshold value. Because of which the system has classified that the person is feeling drowsy.



Figure: Drowsy Output 3

In the above output drowsiness is detected because the person has Yawned, and closed the eyes for too long, so the EAR and MAR values falls below the threshold values. Because of which the system has classified that the person is feeling drowsy.

2) Frames Classified as Active (not drowsy)

Here the input will be the continuous stream of video, if the driver seems to be detected as active then EAR and MAR values will be continuously tracked.

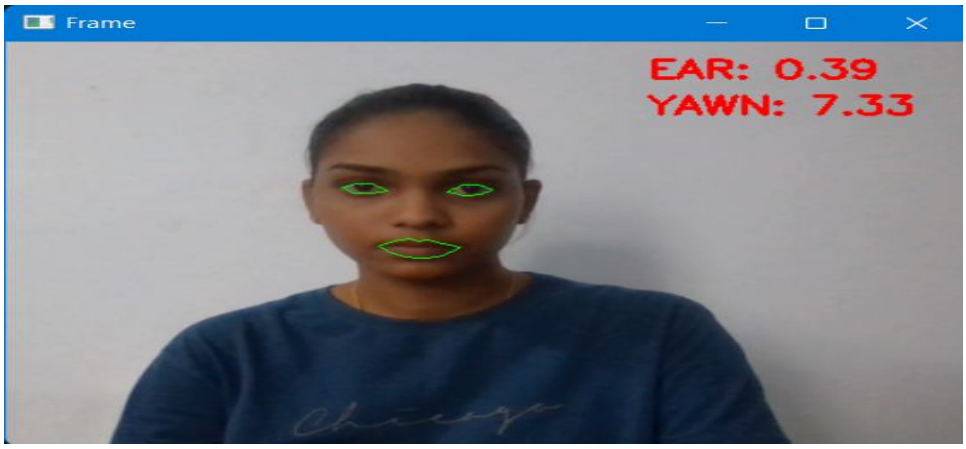


Figure: Active Output 1

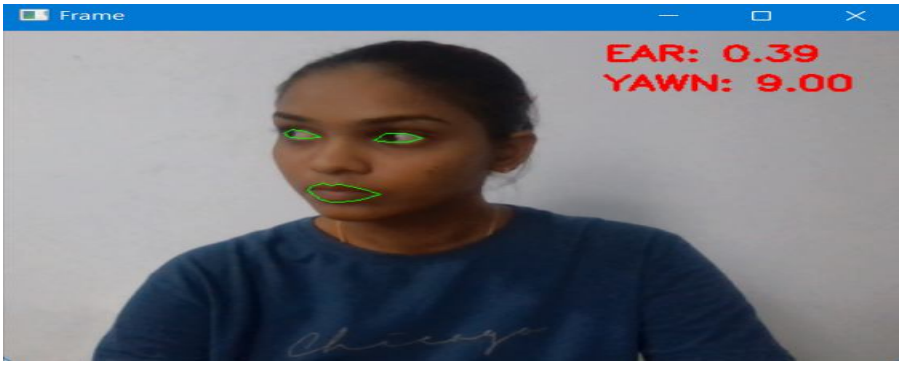


Figure: Active Output 2

In the above output drowsiness is not detected because the persons EAR and MAR values are in the limits of the respected threshold values, because of which the system has classified that the person is Active.

V. CONCLUSION

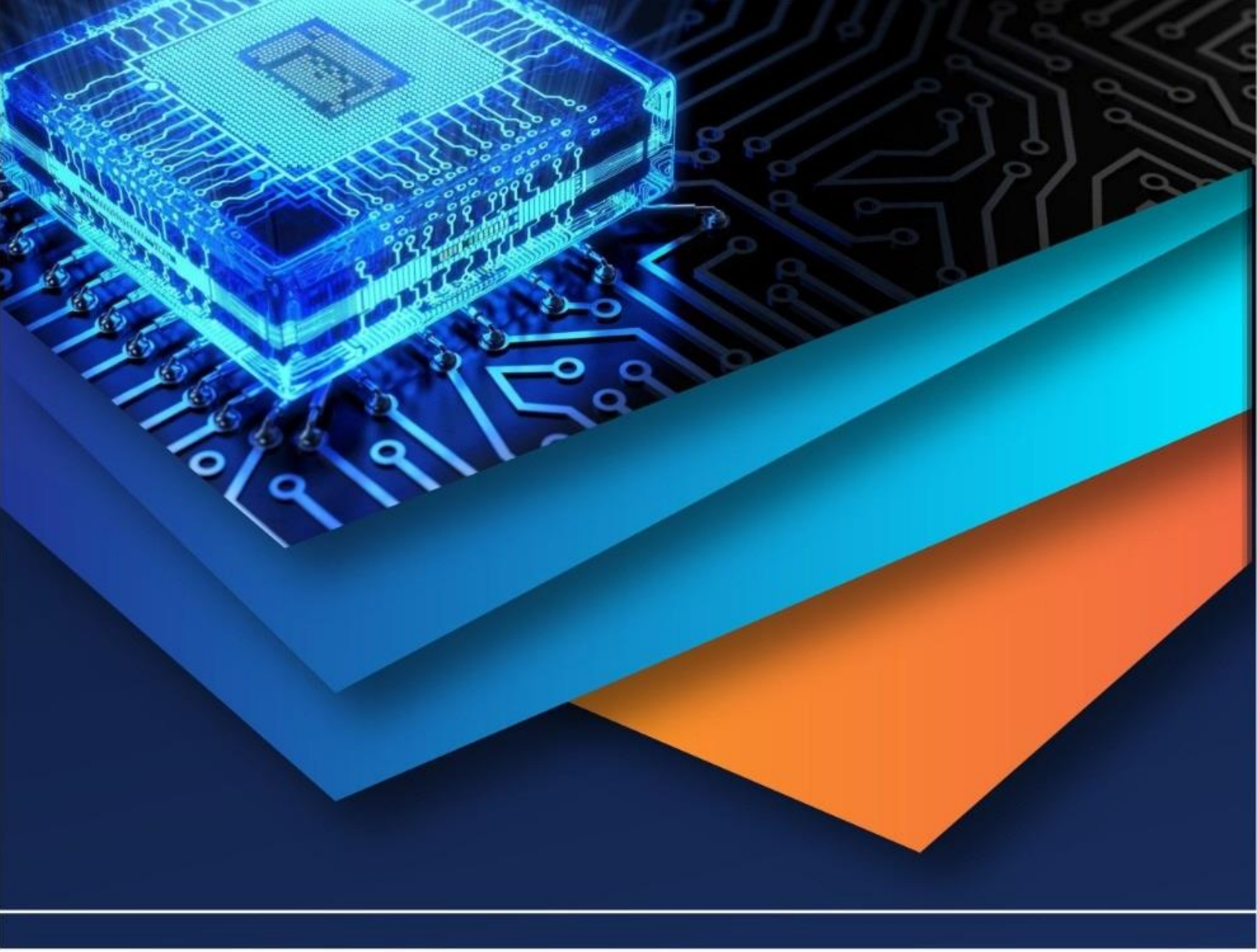
A driver drowsiness detection system using OpenCV and CNN is a promising technology that has the potential to improve road safety by alerting drivers when they are getting drowsy or distracted. The system works by analyzing the driver's face and eyes to detect signs of drowsiness, such as drooping eyelids and yawning. The drowsiness detection system can be implemented in every vehicle such that we can prevent road accidents and decrease the death ratio which are caused due to drowsiness. As AI techniques are growing vastly, we can make systems more intelligent to understand the requirements of the hour. We can introduce various models and use different types of algorithms to get the best results. Based on the result analysis of the proposed system, it is concluded that it is effective in detecting drowsiness accurately. The proposed methodology has achieved 99% accuracy. Overall, a driver drowsiness detection system using OpenCV and CNN has the potential to be an effective tool for enhancing road safety and reducing the risk of accidents caused by driver drowsiness.

VI. ACKNOWLEDGMENT

We want to express our deep-felt gratitude and sincere thanks to our guide Mrs. Shilpa Shesham, Assistant Professor, Department of AI, Anurag University, for her skilful guidance, timely suggestions, and encouragement in completing this project. We want to express our profound gratitude to all for having helped us in achieving this dissertation. Finally, we would like to express our heartfelt thanks to our parents, who were very financially and mentally supportive and for their encouragement to achieve our goals.

REFERENCES

[1] Altameem, A. Kumar, R. C. Poonia, S. Kumar and A. K. J. Saudagar, “Early Identification and Detection of Driver Drowsiness by Hybrid Machine Learning”, IEEE Access, Vol. No. 9, 2021.
[2] B. K. Savaş and Y. Becerikli, “Real Time Driver Fatigue Detection System Based on Multi-Task ConNN”, IEEE Access, Vol. No. 8, 2020.
[3] A. Rajkar, N. Kulkarni and A. Raut, ”Driver Drowsiness Detection Using Deep Learning”, ICCET Advances in Intelligent Systems and Computing, Springer, Vol. No. 1354, 2021.
[4] M. J. Flores, J. M. Armingol and A. de la Escalera,“Real-Time Warning System for Driver Drowsiness Detection Using Visual A Information”, Journal of Intelligent and Robotic Systems, Springer, 2019.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)