Driver Drowsiness Detection Using Convolutional Neural Networks

Nehal Mahajan, Yuvraj Singh, Monika
Department of Computer Science and Engineering, Netaji Subhas University of Technology (NSUT), New Delhi, India
nehalmahajan28@gmail.com, yuvrajishaan@gmail.com, monika.phd21@nsut.ac.in

***Abstract*—Driver drowsiness can be caused by various reasons such as sleep deprivation, long hours of driving, monotonous roads such as highways, alcohol and substance use, time of day, and poor work-life balance. This can result in fatal accidents, as drowsy drivers have reduced reaction times, impaired judgment, and are more likely to fall asleep, leading to collisions with other vehicles, pedestrians, or roadside obstacles. Such accidents often occur at high speeds and can be devastating, causing severe injuries or death to both the driver and others involved. Drivers can avoid these accidents if they are warned in time.**

There are various methods for detecting drowsiness, and this paper proposes a deep learning approach using convolutional neural networks to detect drowsiness in drivers by analyzing their face and eye regions.

I. Introduction

Driver drowsiness is a state of extreme fatigue or sleepiness experienced by a driver, due to which they are unable to stay alert, cannot focus on the road, and are unable to react effectively to driving conditions they face on the road. This is usually caused due to lack of sleep, long hours of driving, and monotonous driving on roads like highways, as well as inadequate breaks during long trips. Also, use of alcohol, drugs, and sedatives, heavy use of medications which affect sleep quality also contribute to drivers feeling sleepy. Work-life imbalance like irregular and hectic work schedules, shift work, driving during the body's natural sleep cycle (late night or early morning), and mental fatigue caused due to unrealistic amounts of workload can lead to driver feeling fatigued [3]. Many employees returning after work have secretion of melatonin after long shifts of work. Melatonin is a hormone that signals the body to prepare for sleep, making drivers feel more tired and less alert. It increases naturally during nighttime or in low-light conditions, leading to drowsiness. Not paying attention to health factors like dehydration, poor diet, stress, and untreated sleep disorders like sleep apnea, narcolepsy, or restless leg syndrome are also major reasons for driver drowsiness.

Drowsy driving can have serious consequences which can be life-threatening. Drivers may take longer amounts of time to respond to sudden blockages on the road, which can cause delays in braking or steering. If the driver is not in the right state of mind due to medication or substance use, this can lead to impaired decision-making, increasing the probability of accidents. This also reduces the ability to focus on the road. Drowsy drivers are more likely to cause collisions due to missed lights, not following speed limits, and lack of control of the vehicle. They may shift to the wrong lanes on the road. Extreme fatigue and lack of attention can lead to involuntary periods of sleep lasting 15–20 seconds, which are also a major cause of accidents when truck drivers fall asleep driving at night. Straight stretches of road with no visual stimulation reduce mental engagement, which makes it harder to focus while driving. A drowsy driver is unable to remember the last few stretches they have driven. This makes it difficult for the driver to stay focused or respond to new situations. This cognitive decline can put the driver and also others on the road at risk of fatal accidents, especially in situations where heightened awareness and rapid actions are required.

From January 2012 to March 2023, the Yamuna Expressway Industrial Development Authority reported that over 44% of accidents on the 165.5 km-long, six-lane Yamuna Expressway were caused by drivers falling asleep at the wheel. Additionally, around 18% of the accidents during this period were attributed to overspeeding. In 2015, the Ministry of Road Transport and Highways (MoRTH) documented 3,081 accidents resulting from driver fatigue, which led to 706 deaths and 3,383 injuries. Highway patrol authorities estimate that sleep-deprived drivers are involved in roughly 40% of road accidents [4]. Also, these numbers might not be accurate as it is really difficult to predict if the accident was caused due to driver drowsiness and lack of sleep. This shows the urgent need for enhanced safety strategies, including better tracking of fatigue-related incidents and measures to mitigate drowsy driving.

Due to advancements in technology and developments in the field of Artificial Intelligence, it is possible to detect if the driver is drowsy and also alert them, which can prevent the occurrence of road accidents. This includes leveraging a combination of various techniques, such as computer vision, machine learning, and physiological monitoring, which helps assess the driver's alertness in real time. A key indicator of drowsiness is facial expressions; facial movements like eye blinking rate, eyelid closure, and head positioning help identify signs of fatigue. AI can recognize changes in mouth shape or size which indicates yawning. Algorithms can also be used to detect partially closed eyes, which helps assess alertness. Drowsy drivers can also drift their gaze from the road, which can also be detected using such systems. This can be implemented using computer vision-based systems to identify specific facial features and movement patterns and deep learning models for facial recognition like convolutional

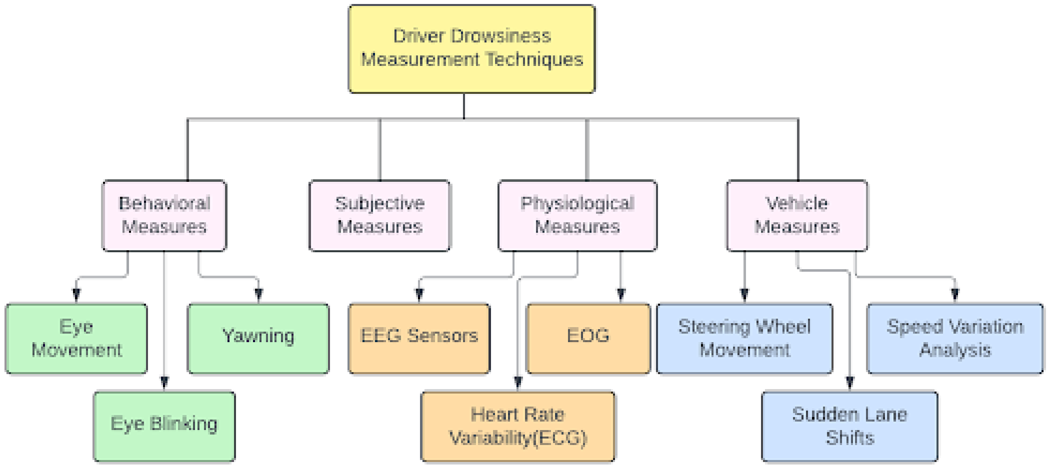

Fig. 1: Driver Drowsiness Measurement Techniques

neural networks (CNNs). The system analyzes these patterns and, once a threshold is reached, it can trigger an alert, helping to keep the driver awake and aware.

In addition to facial recognition, some systems even monitor driving patterns like sudden lane shifts, irregular steering, or inconsistent speed, which may indicate driver drowsiness. AI systems that use cameras and sensors can detect when the vehicle crosses lane markings. Models can analyze steering patterns and detect erratic or delayed steering inputs. AI can process data from speed sensors to identify unusual speed variations.

Beyond visual and driving pattern recognition, advanced AI-based systems are using biometric sensors that track physiological data such as heart rate variability, skin temperature, or brain activity to detect fatigue through wearable devices or built-in car sensors. This can help gain deeper insights into the driver's level of alertness. Heart rate tends to decrease when the driver is drowsy, and increased fatigue leads to lower skin conductance, which can be detected using wearable devices or sensors in the seat of the vehicle. Some advanced AI systems also use electroencephalography (EEG) sensors [5] integrated into headbands and caps which monitor brain activity and can help detect fatigue.

When signs of drowsiness are detected using these AI systems, they can release immediate warnings through sounds, vibrations, alarms, visual alerts like dashboard notifications, or physical interventions like vibrations on the seat of the vehicle or resistance on the steering wheel. During long periods of driving, the AI systems may recommend rest breaks or take control as well in semi-autonomous vehicles. These warnings alert the driver to rest or take necessary precautions, thus significantly reducing the risk of accidents caused by driver fatigue.

II. Motivation of the Work

Drowsy driving is a significant cause of accidents globally, responsible for countless injuries, fatalities, and substantial economic costs each year. With fatigue impairing driver alertness and response times, studies show that a high percentage of road accidents are directly linked to drowsiness. Current detection methods, however, often depend on complex, multi-factor systems that are difficult to implement for real-time use in typical driving conditions. This research aims to address this gap by focusing on two primary indicators—eye movement and yawning—enabling a more streamlined, efficient, and accurate approach. Utilizing deep learning, this simplified system can continuously monitor drowsiness in real time, potentially preventing accidents before they occur, thus enhancing overall road safety and reducing the heavy toll of drowsy driving.

III. Literature Review

Numerous studies have been conducted to address the critical issue of driver drowsiness detection, a few of which are reviewed below.

Gao et al. (2019) [18] proposed a recurrence network-based convolutional neural network (RN-CNN) model to detect driver fatigue using electroencephalogram (EEG) signals. Their approach involves constructing a multiplex recurrence network (RN) from EEG data, followed by the application of a convolutional neural network (CNN) to extract and classify features. The model achieved an impressive mean accuracy of 92.95%, highlighting its effectiveness. Future research might focus on integrating this method into real-time systems and improving its adaptability for practical applications across various driving conditions.

De Naurois et al. (2018) [19] explored the difficulties in monitoring driver drowsiness, particularly due to significant individual differences that impact detection accuracy. To enhance performance, they trained artificial neural networks (ANNs) using data collected from 20 drivers, developing two models:

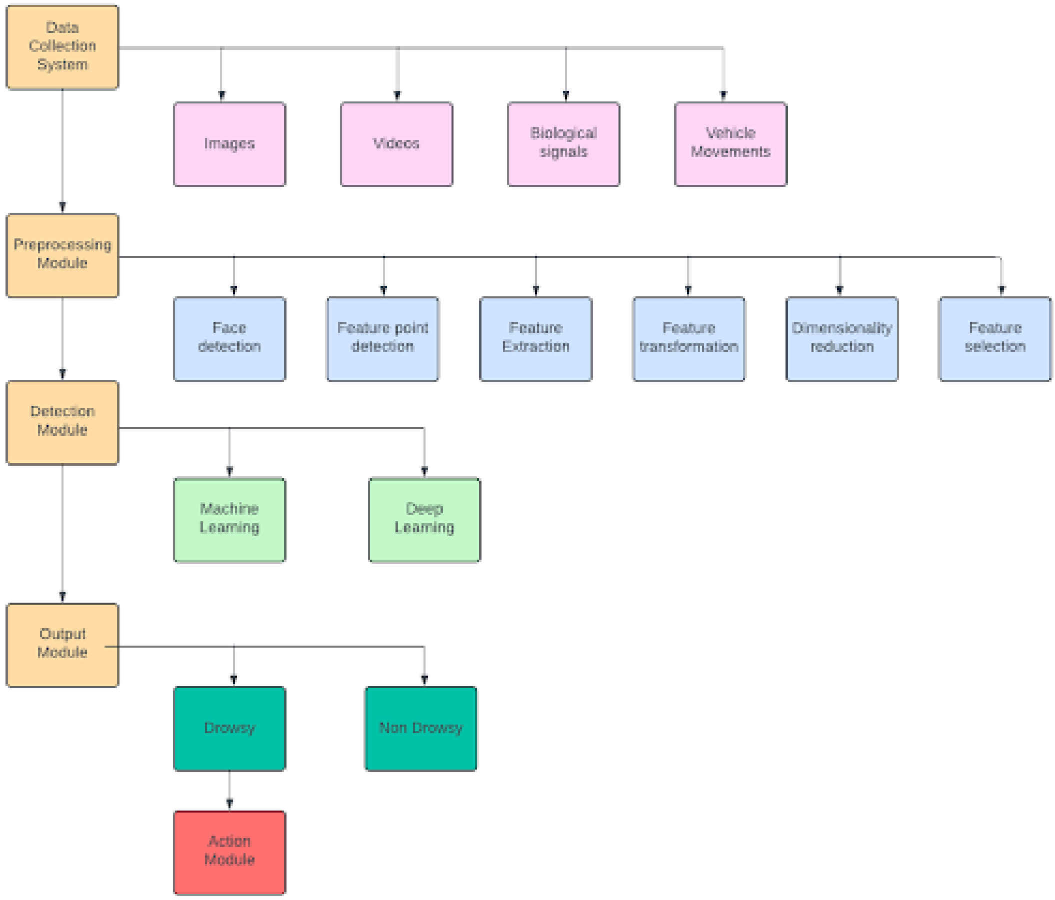

Fig. 2: General Flow in Driver Drowsiness Detection Systems

one for detecting drowsiness on a minute-by-minute basis and another for predicting the time until moderate drowsiness onset. The adapted models demonstrated approximately 40% improvement in prediction accuracy and an 80% increase in detection rates. Future studies could refine these models for real-time use and enhance their adaptability to different driving scenarios.

Zhang et al. (2020) [20] presented a method for detecting driver drowsiness by identifying facial landmarks within video frames. They utilized a residual-based deep 3D convolutional neural network (CNN) to classify sequences of facial data and produce drowsiness probability values, which were then processed by a recurrent neural network. The approach achieved an accuracy of 88.6% on a public dataset, demonstrating the effectiveness of 3D CNNs in capturing spatiotemporal features. Future research could aim to enhance real-time applicability and robustness under diverse driving conditions.

Dua et al. (2020) [21] addressed the issue of drowsy driving, which affects approximately 20% of drivers according to the National Sleep Foundation. They developed a detection system utilizing four deep learning models—AlexNet, VGG-FaceNet, FlowImageNet, and ResNet—to analyze RGB videos of drivers. By assessing features such as hand gestures and facial expressions, the models classified driver states into non-drowsiness, drowsiness with eye blinking, yawning, and nodding. The system achieved an accuracy of 85% using an ensemble algorithm and a SoftMax classifier. Future work could concentrate on enhancing model accuracy and adaptability for real-time monitoring in various driving environments.

Siddiqui et al. (2021) [22] proposed a method to classify drowsy and non-drowsy driver states using non-invasive impulsive radio ultra-wideband (IR-UWB) radar to detect respiration rates. Data collected from 40 subjects were used to train several machine learning models, including support vector machines (SVM), decision trees, logistic regression, and multilayer perceptrons. Among these, the SVM model achieved the highest accuracy of 87% and an F1-score of 73%. Future research could focus on integrating real-time data monitoring to enhance driver drowsiness detection and improve overall road safety.

Gwak et al. (2020) [23] investigated driver alertness states using a driving simulator to measure physiological signals and behavioral indices associated with drowsiness. They applied various machine learning models—including support vector machines (SVM), k-nearest neighbors (KNN), and random forest (RF)—with an ensemble algorithm achieving an accu-

TABLE I: Literature Reviews

S.NO	Authors	Classification Methods	Drowsiness Measures	Dataset	Performance
1	Guo et al. (2016) [6]	Bayesian Network (BN)	Physiological and behavioral signals: heart/pulse rates, eyelid and gaze movements, head posture	Simulator study with 21 participants	Accuracy: 79.50%
2	Bakheet and Al Hamadi (2021) [7]	HOG Features	Analysis of driver images, focusing on eye regions	NTHU-DDD dataset	Accuracy: 85.62%
3	Chen et al. (2021) [8]	LSTM and CNN	Facial analysis, concentrating on eyes and face area	NTHU-DDD dataset	Accuracy: 93.30%
4	Guo and Markoni (2018) [9]	Hybrid CNN-LSTM	Detection of facial landmarks (eyes and mouth)	Public drowsy driver dataset	Accuracy: 84.85%
5	Wijnands et al. (2019) [10]	3D Neural Networks	Behavioral cues: yawning, nodding, looking aside, talking, laughing, eye closure	DDD dataset	Accuracy: 80.8%
6	Jabbar et al. (2020) [12]	CNN	Indicators like yawning, slow blinking, and lethargic head movements	NTHU-DDD dataset	Accuracy: 88.00%
7	Deng W, Wu R (2019) [13]	Multiple CNN-kernelized correlation filters	Mouth and eye	CelebA, YawDD datasets	Accuracy: 92.00%
8	Moujahid A et al. (2021) [15]	SVM	Eye, head, mouth	NTHU-DDD dataset	Accuracy: 79.84%
9	Celecia A et al. (2020) [16]	Mamdani fuzzy inference system	Eye and mouth	300-W dataset	Accuracy: 95.5%
10	Maior C.B.S. et al. (2020) [17]	MLP, RF, SVM	Eye	DROZY dataset	SVM: 94.9%

racy of 95.4% in distinguishing between alert and moderately drowsy states. These results demonstrate the potential for effective early detection systems for driver drowsiness. Future research could focus on refining these models for real-time applications and testing across a wider range of drivers to improve reliability.

Yu et al. (2019) [24] introduced a condition-adaptive framework for detecting driver drowsiness using a 3D deep convolutional neural network (3D-DCNN). The framework employs spatio-temporal representation learning and feature fusion to classify driving conditions, such as the presence of glasses and facial movements, resulting in a more discriminative representation. Evaluation on the NTHU Drowsy Driver Detection dataset achieved an accuracy of 76.20% and an F1-score of 76.50%. Future research may aim to improve the framework's accuracy and applicability in real-time across diverse driving environments.

IV. Problem Statement

Driver drowsiness poses a serious threat to road safety, significantly increasing the likelihood of accidents by slowing reaction times, diminishing concentration, and impairing judgment. Fatigued drivers often miss crucial road signals, drift from their lanes, or even fall asleep momentarily—all of which can result in severe accidents, particularly at high speeds. In India, this issue is especially critical; data from the Ministry of Road Transport and Highways indicates that fatigue among drivers contributes to more than 40% of accidents on highways [25]. These incidents frequently lead to serious injuries and fatalities, highlighting the pressing need for effective methods to identify and mitigate driver drowsiness.

The gravity of this problem is further underscored by studies showing that drowsy drivers exhibit impairments comparable to those of individuals under the influence of alcohol. Research indicates that they experience similar reductions in reaction times and decision-making abilities. In the United States alone, drowsy driving is associated with approximately 100,000 reported crashes annually [3], leading to significant loss of life and long-term injuries. Addressing this issue is vital not only for reducing preventable accidents but also for enhancing overall road safety.

Despite the existence of various drowsiness detection methods, many rely on multiple visual and non-visual indicators, requiring complex and costly setups. Techniques that monitor physiological signals—such as heart rate or electroencephalogram (EEG) [5]—or assess vehicle behavior through specialized sensors can be invasive and challenging to implement on a broad scale. Integrating multiple sensors and hardware adds complexity, making widespread adoption difficult for drivers and fleet operators.

This study proposes a streamlined approach focusing on eye closure and yawning, two easily observable and effective visual indicators of drowsiness. By utilizing a Convolutional Neural Network (CNN) [26] to classify images as either eyes open or closed, and yawning or not yawning, this method offers a cost-effective and non-intrusive alternative that requires only standard camera input. This simplification reduces the technological barriers associated with traditional drowsiness detection methods, enhancing the practicality and accessibility of implementation in vehicles. The aim is to contribute significantly to the development of effective drowsiness detection systems by concentrating on these visual factors, thereby addressing the critical issue of driver fatigue and its associated risks.

V. Datasets

There exist several standardized datasets that researchers have used for drowsiness detection, as summarized in Table II.

This paper utilizes two of these datasets. The first is the YawDD Video dataset [2], which comprises videos recorded by a camera mounted on a car dashboard. This dataset includes male and female drivers, some wearing glasses and others without. It features a collection of videos captured by an in-car camera, showcasing real drivers engaged in various activities such as talking, singing, yawning, and more. The drivers represent diverse facial features, including different ethnicities and the presence of glasses or sunglasses. The second dataset used is the Closed Eyes in the Wild (CEW) dataset [1]. This dataset contains 2,423 subjects, with 1,192 people having closed eyes and 1,231 people with open eyes.

VI. Methodology

A. Data Preprocessing

Data augmentation [38] is a process that improves the generalization of a model by creating variations of the original image data. In our case, several data augmentation techniques are implemented. Random cropping is used to randomly crop a section of the image and resize it to the original size of the image. This technique helps the model learn features from different parts of the image, making it more robust to variations in the position of the eyes or the face in the image.

Horizontal flipping is used to flip the image horizontally. This technique helps the model learn features from different orientations and hence makes it more robust to variations in the orientation of the head or face in the image. By using data augmentation, the model can learn to identify drowsiness under various conditions, such as different lighting conditions, different head and eye positions, and different driver appearances.

To enhance the model's robustness, additional data augmentations such as zooming ($\pm 20\%$), rotation (up to $30°$), and horizontal flipping are applied using TensorFlow. This ensures the model can adapt to various lighting conditions, head angles, and individual driver traits.

1) Data Transformation

The data transformation preprocesses images to ensure uniformity. Images are loaded from structured directories containing categories like "yawn," "no_yawn," "closed eye," and "open eye." For yawning detection, Haar cascade classifiers [14] are utilized to detect facial regions. These classifiers work by applying a sliding window approach and evaluating regions against pre-trained Haar features, isolating facial areas of interest. Detected faces are cropped and resized to 145×145 pixels using OpenCV's resizing functions.

For eye data, images undergo direct resizing to the same dimensions, maintaining consistency across datasets. The preprocessed facial and eye data are then combined into a unified array, enabling analysis of yawning and eye closure patterns.

2) Data Splitting and Normalization

The processed data is split into training and testing sets, with a 70–30 split. Normalization is performed by scaling pixel values to the $[0, 1]$ range, which helps in achieving faster and more stable convergence during model training by ensuring that the input features are on a similar scale.

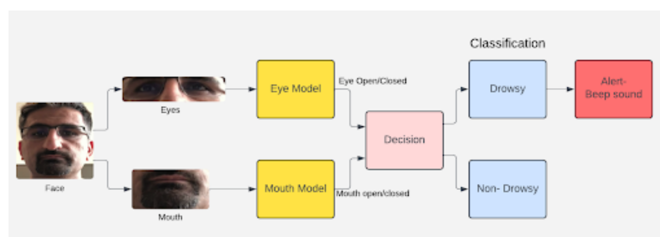

Fig. 3: Workflow of the Driver Drowsiness Detection System

B. Model Description

The Convolutional Neural Network (CNN) [26] used for driver drowsiness detection is designed for efficient and scalable image classification. The input layer processes 4D tensors with dimensions corresponding to the batch size, image height and width (145×145 pixels), and color channels [11] (RGB). The architecture begins with a convolutional layer featuring 512 filters and a 3×3 kernel, employing the ReLU activation function [36]. This operation extracts local features, such as edges and textures, crucial for identifying drowsiness-related cues. Each convolutional layer is followed by max-pooling layers with a 2×2 pool size, reducing spatial dimensions and minimizing computational overhead while preserving essential features.

The network deepens progressively, employing additional convolutional layers with 512 and 256 filters, which capture increasingly abstract patterns. The hierarchical feature extraction culminates in a Flatten layer [37], converting the 3D feature maps into a 1D vector suitable for dense layers. This transformation bridges the convolutional layers with the fully

TABLE II: Summary of Drowsiness Detection Datasets

Ref./Year	**Dataset Name**	**Description**
[2], 2020	YawDD	Videos of drivers showing drowsiness signs, captured from two camera angles in various scenarios.
[27], 2016	DROZY	Contains synchronized data from 14 individuals, including EEG and eye-tracking information.
[1], 2014	CEW	Includes 2,423 subjects' eye images, focusing on open and closed eye detection.
[28], 2016	NTHU-DDD	Contains 36 subjects in various driving scenarios, focusing on drowsiness indicators like yawning.
[29], 2015	CelebA	A large dataset of 200K celebrity images with 40 binary attribute labels per image.
[30], 2018	MRL Eye	Comprises 84,898 images depicting various eye states and lighting conditions from 37 individuals.
[31], 2020	ZJU Gallery	Multi-view dataset of dynamic human videos recorded under different lighting conditions.
[32], 2013	300-W	Includes 600 images (300 indoor and 300 outdoor) with 68 facial landmark annotations.
[33], 1999	MIT Polysomnographic Database	Contains polysomnographic recordings for sleep studies, focusing on sleep disorders such as sleep apnea.
[34], 2000	Sleep-EDF Database	Offers whole-night sleep recordings with EEG and other physiological signals for sleep research.
[35], 2019	UTA Real-Life Drowsiness Dataset (UTA-RLDD)	Provides 30 hours of RGB video footage of 60 subjects, labeled for alertness, low vigilance, and drowsiness.

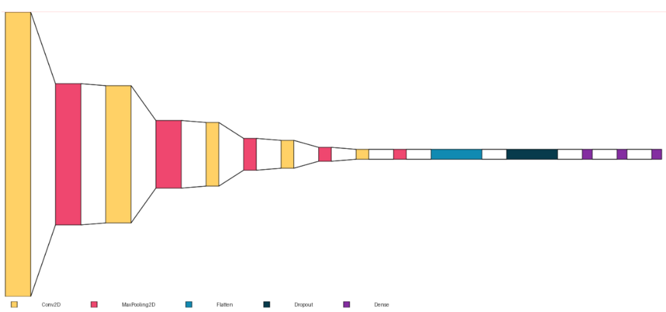

Fig. 4: Network Architecture Diagram

connected layers, where high-level relationships in the data are modeled.

To prevent overfitting, a Dropout layer [42] with a dropout rate of 50% is integrated before the dense layers. Dropout stochastically disables neurons during training, preventing over-reliance on specific pathways and encouraging the model to generalize better. The fully connected section includes two dense layers with 128 and 64 neurons, respectively, both using the ReLU activation function to capture non-linear dependencies. The final output layer uses a softmax activation function [39], producing a probability distribution over the four classes: yawn, no yawn, closed eyes, and open eyes.

The model is compiled with the categorical cross-entropy loss function [40], optimal for multi-class classification tasks, and the Adam optimizer [41], which dynamically adjusts learning rates for faster convergence. Training incorporates mini-batch gradient descent, processing batches of augmented data. Additionally, the model architecture, implemented using TensorFlow's Sequential API, is designed to allow seamless backpropagation and optimization. This CNN architecture, designed for efficiency in training and evaluation, is well-suited to the task of detecting driver drowsiness from images, providing a robust framework for real-time image classification.

VII. Results

The network is trained on an augmented dataset over 30 epochs, achieving a training accuracy of 96.21% and a testing accuracy of 96.54%. To assess the model's performance, training and testing loss and accuracy are plotted using Matplotlib, as shown in Fig. 5 and Fig. 6.

Fig. 5: Training and Testing Accuracy

The model exhibits excellent precision and recall for all four classes, showing especially strong results in detecting Closed and Open eye labels, with F1-scores of 0.95 and 0.96, respectively. The weighted average F1-score is 0.94, while the macro average F1-score is 0.92, indicating the model's strong capability in identifying driver drowsiness.

The weighted average Recall (Sensitivity) is 94.10% and the weighted average Specificity comes out to be 97.95%.

The results indicate that the developed system has the potential to be used in real-world applications, such as in

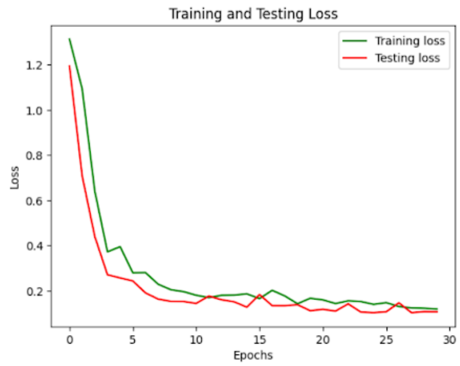

Fig. 6: Training and Testing Loss

vehicles equipped with a camera to monitor the driver's state. The system can provide an alert to the driver in case of drowsiness, which can help prevent accidents and improve road safety.

```
              precision    recall  f1-score

        yawn       0.88      0.84      0.86
     no_yawn       0.87      0.97      0.92
      Closed       0.98      0.93      0.95
        Open       0.96      0.97      0.96

    accuracy                           0.94
   macro avg       0.92      0.93      0.92
weighted avg       0.94      0.94      0.94


Specificity for each class:
yawn: 0.9864
no_yawn: 0.9782
Closed: 0.9862
Open: 0.9716

Macro Average Specificity: 0.9806
Weighted Average Specificity: 0.9795
```

Fig. 7: Performance Characteristics

VIII. Comparison with Base Paper

TABLE III: Comparison with Base Paper

	Base Paper [44]	**Proposed Paper**
Methodology	MTCNN	Haar Cascade and CNN
Dataset	IT Company - Biteda	YAWDD and CEW
Accuracy	93.623%	96.54%
Sensitivity	93.643%	94.10%
Specificity	60.882%	97.95%
Parameters	Eye and Mouth	Eye and Mouth

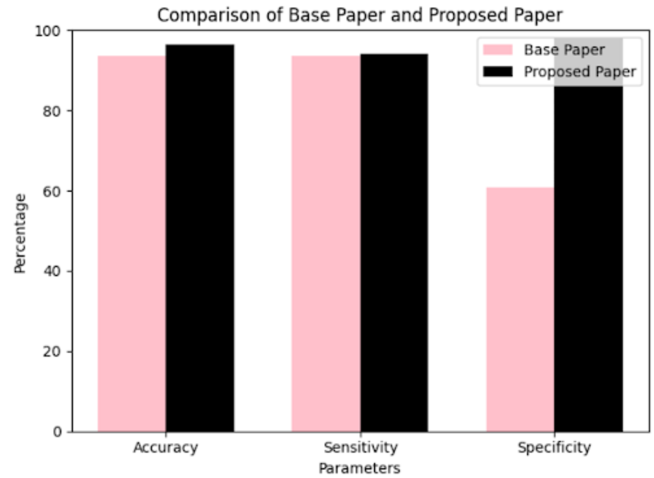

Fig. 8: Comparison of Parameters in Base Paper and Proposed Paper

IX. Hardware and Software Requirements

A. Hardware Requirements

1) GPU: Two NVIDIA T4 units on Kaggle for efficient training and inference.
2) CPU: Standard CPU for local execution; longer processing times without GPU.
3) RAM: Minimum of 16 GB for effective data processing and training.
4) Storage: At least 50 GB for datasets, model files, and outputs.

B. Software Requirements

1) Operating System:
   a) Windows for local development.
2) Cloud Environment:
   a) Kaggle Notebook for cloud execution.
3) Python:
   a) Version 3.6 or higher.
4) Libraries:
   a) TensorFlow: Model building and training.
   b) Keras: High-level API for model construction.
   c) OpenCV: Image processing, including face detection.
   d) NumPy: Numerical computations and image data handling.
   e) scikit-learn: Data preprocessing and encoding.
   f) Matplotlib: Visualizing training progress.

X. Future Work and Conclusion

The driver drowsiness detection system was successfully developed using a Convolutional Neural Network (CNN) architecture, achieving a maximum testing accuracy of 96.54% on the YawDD and CEW datasets. These datasets contributed to the robustness of the model, as YawDD includes varying levels of drowsiness and CEW provides a range of emotional expressions, enabling the system to perform effectively across different driver states. This research highlights the potential of deploying CNN-based drowsiness detection in real-world applications, where vehicles equipped with cameras could use the system to monitor drivers and issue alerts when drowsiness is detected, ultimately helping to reduce accidents and enhance road safety.

Future research in deep learning-based driver drowsiness detection systems can expand on these findings through several key directions. First, integrating multimodal data, such as combining facial cues, eye gaze tracking, head movement analysis, and physiological signals (e.g., heart rate, brain activity), could improve the system's depth of understanding. Advanced fusion techniques, such as multi-stream networks or attention mechanisms, can be used to selectively attend to the most relevant signals, creating a more holistic representation of drowsiness.

Improving model generalization across diverse populations and environmental conditions remains essential for effective deployment. Current models, often trained on specific datasets, may struggle to perform reliably in real-world scenarios with variations in demographics, lighting, and camera angles. Future work could address these challenges through domain adaptation, data augmentation, and adversarial training, which would help create models that are robust under varied conditions.

Efficient, real-time implementation is another critical focus. Practical deployment requires optimization techniques like model compression, quantization, and the use of hardware accelerators to reduce computational demands. Developing lightweight architectures specifically designed for drowsiness detection can further ensure low-latency processing, allowing for quick alerts and maximizing safety on the road.

Finally, transfer learning offers a promising approach to overcoming the limitation of annotated drowsiness datasets. By pre-training on larger, general datasets and then fine-tuning on domain-specific data, models can leverage learned features to improve performance, even with limited labeled samples. This approach can enhance robustness and generalizability, addressing data scarcity issues.

By advancing these aspects—multi-modal integration, improved generalization, efficient real-time processing, and data-efficient transfer learning—future research can drive significant improvements in driver drowsiness detection systems. These enhancements will lead to more practical, reliable, and accessible systems, further contributing to road safety and reducing the risks associated with drowsy driving.

References

[1] Song F., Tan X., Liu X., Chen S. (2014). Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. *Pattern Recognition*, 47(9), 2825–2838.

[2] Abtahi S., Omidyeganeh M., Shirmohammadi S., Hariri B. (2014). YawDD: a yawning detection dataset. In *Proceedings of the 5th ACM Multimedia Systems Conference*, pp. 24–28.

[3] UCLA Health. (n.d.). Drowsy driving.

[4] Times of India. (2023). No data on driver fatigue-related accidents reveals RTI reply.

[5] NHS. (n.d.). Electroencephalogram (EEG).

[6] Guo W., Zhang B., Xia L., Shi S., Zhang X., She J. (2016). Driver drowsiness detection model identification with Bayesian network structure learning method. In *Chinese Control and Decision Conference (CCDC) 2016*, pp. 131–136.

[7] Bakheet S., Al-Hamadi A. (2021). A Framework for Instantaneous Driver Drowsiness Detection Based on Improved HOG Features and Naïve Bayesian Classification. *Brain Sciences*, 11(2), 240.

[8] Chen S., Wang Z., Chen W. (2021). Driver Drowsiness Estimation Based on Factorized Bilinear Feature Fusion and a Long-Short-Term Recurrent Convolutional Network. *Information*, 12(3), 94.

[9] Guo J.-M., Markoni H. (2019). Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. *Multimedia Tools and Applications*, 78(20), 29059–29087.

[10] Wijnands J.S., Thompson J., Nice K.A., Aschwanden G.D.P.A., Stevenson M. (2020). Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. *Neural Computing and Applications*, 32(13), 9731–9743.

[11] Medium. (n.d.). Understanding the Concept of Channels in an Image.

[12] Jabbar R., Shinoy M., Kharbeche M., Al-Khalifa K., Krichen M., Barkaoui K. (2020). Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application. In *International Conference on Internet of Things (ICIoT) 2020*, Doha, Qatar.

[13] Deng W., Wu R. (2019). Real-time driver-drowsiness detection system using facial features. *IEEE Access*, 7, 118727–118738.

[14] Medium. (n.d.). Haar Cascades Explained.

[15] Moujahid A., Dornaika F., Arganda-Carreras I., Reta J. (2021). Efficient and compact face descriptor for driver drowsiness detection. *Expert Systems with Applications*, 168, 114334.

[16] Celecia A., Figueiredo K., Vellasco M., González R. (2020). A portable fuzzy driver drowsiness estimation system. *Sensors*, 20(15), 4093.

[17] Maior C.B.S., das Chagas Moura M.J., Santana J.M.M., Lins I.D. (2020). Real-time classification for autonomous drowsiness detection using eye aspect ratio. *Expert Systems with Applications*, 158, 113505.

[18] Gao Z.-K., Li Y.-L., Yang Y.-X., Ma C. (2019). A recurrence network-based convolutional neural network for fatigue driving detection from EEG. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(11), 113125.

[19] De Naurois C.J., Bourdin C., Bougard C., Vercher J.-L. (2018). Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. *Accident Analysis & Prevention*, 121, 118–128.

[20] Zhang X., Wang X., Yang X., Xu C., Zhu X., Wei J. (2020). Driver Drowsiness Detection Using Mixed-effect Ordered Logit Model Considering Time Cumulative Effect. *Analytic Methods in Accident Research*, 26, 100114.

[21] Dua M., Shakshi, Singla R., Raj S., Jangra A. (2021). Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Computing and Applications*, 33(8), 3155–3168.

[22] Siddiqui H.U.R., Saleem A.A., Brown R., Bademci B., Lee E., Rustam F., Dudley S. (2021). Non-Invasive Driver Drowsiness Detection System. *Sensors*, 21(14), 4833.

[23] Gwak J., Hirao A., Shino M. (2020). An Investigation of Early Detection of Driver Drowsiness Using Ensemble Machine Learning Based on Hybrid Sensing. *Applied Sciences*, 10(8), 2890.

[24] Yu J., Park S., Lee S., Jeon M. (2019). Driver drowsiness detection using condition-adaptive representation learning framework. *IEEE Transactions on Intelligent Transportation Systems*, 20(11), 4206–4218.

[25] The Hindu. (2022). Sleep-deprived drivers responsible for 40% of road accidents.

[26] MathWorks. (n.d.). Introduction to deep learning: What are convolutional neural networks.

[27] Massoz Q., Langohr T., François C., Verly J.G. (2016). The ULg multimodality drowsiness database (called DROZY) and examples of use. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–7.

[28] Weng C.H., Lai Y.H., Lai S.H. (2016). Driver drowsiness detection via a hierarchical temporal deep belief network. In *Asian Conference on Computer Vision*, Springer, Cham, pp. 117–133.

[29] Liu Z., Luo P., Wang X., Tang X. (2015). Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3730–3738.

[30] Fusek R. (2018). Pupil localization using geodesic distance. In *International Symposium on Visual Computing*, Springer, Cham, pp. 433–444.

[31] Zhou K., Wang K., Yang K. (2020). A robust monocular depth estimation framework based on light-weight ERF-PSPNet for day-night driving scenes. *Journal of Physics: Conference Series*, 1518(1), 012051.
[32] Sagonas C., Tzimiropoulos G., Zafeiriou S., Pantic M. (2013). 300 Faces In-the-Wild Challenge: The first facial landmark localization challenge. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 397–403.
[33] Ichimaru Y., Moody G.B. (1999). Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences*, 53(2), 175–177.
[34] Kemp B., Zwinderman A.H., Tuk B., Kamphuisen H.A., Oberye J.J. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194.
[35] Ghoddoosian R., Galib M., Athitsos V. (2019). A realistic dataset and baseline temporal model for early drowsiness detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*,.
[36] BuiltIn. (n.d.). ReLU activation function.
[37] Deeplizard. (2019). Exploring the ReLU activation function.
[38] Hoh S. (2020). Exploring data augmentation techniques for image datasets.
[39] Pinecone. (n.d.). Softmax activation function.
[40] DataCamp. (n.d.). The cross-entropy loss function in machine learning.
[41] BuiltIn. (n.d.). Adam optimization.
[42] Keras. (n.d.). Dropout.
[43] BuiltIn. (n.d.). Transfer learning.
[44] ILT In. (n.d.). Adam optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Wiley.