iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.69131

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Novel CNN Based Model Using LSTM for Driver Drowsiness Detection

Tulip Gautam[1], Suhani Sharma[2], Snigdha Mishra[3], Jassica Rajora[4], Khyati Ahlawat[5]
Computer Science Engineering Department, Indira Gandhi Delhi Technical University for Women, Delhi-110006, India

Abstract: Drowsiness affects human cognitive functioning, response time, and decision-making capacity, which is a major risk factor in road safety. Drowsiness-caused accidents are increasingly common, which indicates the need for real-time drowsiness detection systems. The purpose of this study is to detect signs of fatigue through facial characteristics such as eye closure, yawning, and tilting of the head in binary and multi-class classification functions. The study proposes a hybrid model, which integrates Convolution Neural Networks (CNN) and Long Short-Term Memory (LSTM), which uses a camera-based real -time detection system with DLIB for face tracking. CNN extracts spatial functions, while LSTM preserves temporal dependencies and improves the accuracy of the detection. The experimental result model shows a significant improvement in performance, receiving an accuracy of 96.3%, precision rate of 96.45%, recall rate of 96.33%, and F-measure of 96.32%. The proposed model optimizes feature selection, reduces false alarms, and increases reliability, making this a strong and effective solution to the driver's safety monitoring systems.
Keywords: CNN, LSTM, Drowsiness Detection, Eye Closure, Yawning, and Head Tilt Detection.

I. INTRODUCTION

Drowsiness is a significant and widespread issue that affects 17 cognitive functions, attentiveness, and overall productivity. Whether it manifests as fatigue-induced lapses in concentration while driving, a decline in focus during lectures, or reduced efficiency in the workplace, the consequences of drowsiness extend well beyond just feeling tired. Fatigue is a leading contributor to road accidents, accounting for a substantial number of injuries and fatalities each year [1]. Drowsy driving is a major public safety concern since studies show that it can compromise your reaction time and decision-making skills just as much as drunk driving [2]. Given its widespread implications, there is a growing need for intelligent systems that can spot the signs of drowsiness early and step in to help, before things take a turn for the worse.

Conventional methods of detecting drowsiness like self-evaluations, vital sign monitoring, or relying on someone's observations, often fall short in practical settings [3]. Although wearable technology, such as eye-tracking glasses or EEG headbands, can be useful, its high cost and discomfort make it difficult to utilize regularly [4], [5]. This proves the crucial need for a non-invasive, automated system to recognize drowsiness in the presence of visual cues. Facial expressions and other behaviors such as prolonged eye closure, less frequent blinking, yawning, and tilting the head, can tell you how alert someone is [6]. Thanks to advances in deep learning and computer vision, these subtle facial and behavioral signals can be tapped to detect drowsiness with great accuracy, eliminating the need for these expensive and uncomfortable gadgets [7], [8].

This study introduces a reliable approach that combines the advantages of Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks to detect drowsiness more effectively [9]. While CNNs are good at analyzing stationary images, such as determining whether the eyes are open wide, LSTMs are better at processing input sequentially to identify how something changes over time, such as gradual eye closing or frequent yawning. This combined method not only finds early detection of fatigue in the system but also tracks how fatigue develops over time to improve accuracy and remove false positives [10]. The model employs transfer learning, which is essentially the borrowing of knowledge from previously trained networks to further increase efficiency. This means it can perform well even with less labeled data and still adapt to a wide range of real-world conditions [11].

It is more difficult to identify drowsiness in real-world situations than it is in controlled laboratory settings. Detection can be challenging due to a variety of factors, including shifting lighting, varying facial angles, obstacles like masks or spectacles, and even individual differences in how tiredness manifests [12]. To ensure the model is both accurate and reliable, this research examines it across multiple datasets—including publicly available ones and custom-curated data, representing a wide range of demographics and real-world situations. In this sense, the system is more useful and efficient since it is designed to deal with the unpredictability of real-world situations [13].

This paper is divided into structures so that there can be a clear look at the proposed approach. In the beginning section, it speaks about the motivation of the study, as to why real-time drowsiness detection is so important. Section 2 establishes groundwork by adopting a theoretical background, a review of the methods used, and the problems noticed regarding fatigue detection. This paper proceeds to Section 3 where the study offers a literature review reporting previous work's main findings and identifying the gaps that the present research ideally would achieve. The characteristics of the dataset, including model architecture and feature extraction techniques as well as evaluation metrics are covered in Section 4 of the paper about the experimental setup. The results, quantitative measurements, visual analysis, and the patterns to which were subjected are presented in Section 5. In section 6, the paper is finally packed up by summarizing the main findings that also led to possible reforms and binding future research directions. Section 7 offers an extensive list of references, some of which have corrected this study [14]. This research not only addresses the complications of the detection of drowsiness through an integrated deep-learning method but also paves the way to real-world applications with great scopes to enhance the general quality of security, efficiency, and life.

II. LITERATURE REVIEW

In recent years, with important applications of the detection of drowsiness in safety and productivity that have improved in many areas of interest, the detection of drowsiness has attracted a growing amount of attention. A wide selection of machine learning and deep teaching techniques have been shown to find effective symptoms of drowsiness. The current body of the work is then reviewed, a look at the functions, the data sets used, and the boundaries identified in previous studies [15]. This section wants to understand these aspects and present the basis for the construction of a more intelligent and accurate drowsiness detection system.

In this section, it will provide step-by-step details of progress in the region, fill the gaps in relevant approaches, and suggest some methods for its structure that will correct the design. For example, Singh et al. [16] proposed a system for detecting drowsiness depending on facial places, flashing patterns, and eye movement. They used machine learning classification for real-time monitoring but had a problem with different lighting conditions. After similar lines, Kumar et al. [17] also used the Convolutional Neural Network (CNN) to analyze facial features such as facial features and yawns with high accuracy for the availability of many labeled data for training. To prove the ability and challenges of this work, these studies utilize advanced techniques to detect drowsiness.

Table 1: Literature Survey Table

Reference, Year	Technique	Dataset	Limitations	Accuracy
[18], 2022	Simple Linear Clustering (SLIC), CNN, LSTM, Feature Extraction	Images were collected from 10 drivers (5 males and 5 females) over 6 hours of continuous driving	Limited to 10 drivers; lacks diversity in age and driving conditions	97.78%
[19], 2023	Convolutional Neural Networks (CNN), Image Preprocessing, K-Fold Cross-Validation	Kaggle dataset with 1,400 images	Limited dataset size; possible overfitting with CNN models	94.95%
[20], 2021	OpenCV, Dlib, Eye Aspect Ratio (EAR), HaarCascade Algorithm, Histogram of Local Binary Patterns (LBPH)	Yawdd Dataset	Limited to yawn detection; does not account for other drowsiness indicators	91%
[21], 2022	Recurrent Neural Networks (RNN), Image Preprocessing, Eye State Recognition	Custom dataset from 20 drivers with varied environmental conditions	The dataset is not publicly available; the small sample size	97.2%
[22], 2021	Support Vector Machine (SVM), Feature Engineering, K-NN	A public driver-fatigue dataset with 5000+ images	Some bias due to uncontrolled lighting conditions in images	95.5%

[23], 2023	RandomForest Classifier, Sensor Fusion, Feature Engineering	A dataset combining driver physiological and behavioral data	Limited diversity in driver age groups	89.4%
[24], 2020	CNN, Transfer Learning, Pre-trained Models (ResNet50)	A custom dataset of driver facial expressions	Focuses only on facial expressions, may miss other drowsiness cues	98.3%
[25], 2023	CNN, Real-time Drowsiness Detection, Facial Landmarking	Custom real-time video dataset	Real-time processing challenges; require high computational resources	93.6%
[26], 2020	Gradient Boosting, Eye Movement Tracking, Data Augmentation	A dataset with 1,000 images of drowsy and non-drowsy drivers	Eye movement may not always reflect actual drowsiness	92.1%
[27], 2021	LSTM, Driver Behavior Analysis, Time-Series Data	A time-series dataset of 10,000+ driving hours	Lacks enough varied environmental conditions	90.8%
[28], 2023	Hybrid CNN-RNN Eye Blink Detection	Driver fatigue dataset from commercial trucks	Data focused on long-distance drivers; may not be generalized well	96.2%
[29], 2022	Histogram of Oriented Gradients (HOG) (Jarndal), SVM	2,500+ images from varied lighting conditions and road types	Low lighting and bad weather conditions were not accounted for	88.7%
[30], 2021	Multimodal Fusion, CNN, LSTM	Combined image and physiological data from 50 drivers	Data imbalance between drowsy and non-drowsy drivers	94.3%
[31], 2021	Convolutional Neural Networks (CNN), Gaze Detection	A dataset with 3,000+ gaze-related driver images	The dataset focused solely on gaze, ignoring other fatigue symptoms	92.9%
[32], 2023	K-Means Clustering, Driver Sleepiness Score, Visual Cues	A dataset from wearable devices and cameras	Requires wearable devices; not suitable for all vehicles	89.5%
[33], 2021	Deep Convolutional Neural Networks (DCNN), Facial Expression Recognition	1,200+ images of facial expressions during driving	Limited to facial expression data; does not account for physical signs like yawning	98.6%
[34], 2022	Drowsiness Detection using Fuzzy Logic and EEG signals	Custom EEG signal dataset from drivers	Requires specialized EEG equipment; not practical for all vehicles	87.2%
[35], 2023	Attention Mechanism, Deep Learning	Real-time video data from 15 drivers	Real-time processing delays; limited sample size	95.4%
[36], 2020	Convolutional Neural Networks (CNN), Emotion Recognition	A dataset of drivers' emotional responses collected in driving simulators	Does not cover all drowsiness causes; focuses on emotion alone	90.0%

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

The driver drowsiness detection research presents a number of techniques and datasets that have been investigated to address this critical problem. One of the main methods is Convolutional Neural Networks (CNNs) that are capable of recognizing salient facial features like eye closure, yawning, and head pose [37]. Moreover, LSTM networks are famous for analyzing temporal changes due to their capacity for learning temporal patterns in drowsiness behaviors [38]. These models are typically incorporated into hybrid CNN-LSTM models, which are capable of performing accurate detection using both spatial and temporal signals [39].

The above research demonstrates that hybrid CNN-LSTM models are very powerful. Their performance seems to depend on data quality, especially in the capability of detecting complex spatiotemporal patterns. Obtaining a good dataset on which to train is comparatively simple, especially thanks to the variety of datasets that are available. Accuracy rates have been claimed as being very high in research reports; however, such is generally achieved under optimal conditions [40]. In real-world applications, there are multiple factors, including facial geometry, objects in front of the face (e.g., glasses or masks), and environmental distractions, that can interfere with model performance. This complicates the use of such models on hardware with limited computational capability in real-world applications [41]. There is a need to conduct more research to create datasets that mitigate variability in real-world environments. The model itself also needs to be tuned to make its use feasible in real-time environments. This is one such area that is critical and demands more research, being careful to ensure accuracy.The results indicate that the future research must be on generating representative and diverse data sets that closely represent real conditions. Further, enhancing real-time placement models without interfering is another significant research investigation area. Addressing these limitations, the field can move towards developing effective, scalable, and deployable drowsiness detection systems.

III. EXPERIMENTAL FRAMEWORK

This section provides observations of the experimental structure of this study to develop and evaluate the drowsiness detection system. It emphasizes the proposed approach and describes the integration of CNN for spatial convenience and the integration of LSTM networks for temporary analysis [42]. In addition, this section covers the dataset us ed in the study, used on the classification, and the model performance to assess the planned assessment matrix. By explaining each component of the framework, this section provides basic work to understand the function and provides insight into the design and efficiency of the model [43].

A. Proposed Model

This article introduces a new hybrid model that integrates the CNNS and LSTM networks to detect effective driver entry. By combining the extraordinary ability of CNN -s to extract spatial features from facial images with the capacity of LSTM to record temporary mobility, both the proposed model addresses the immediate symptoms of fatigue (such as yawning or closing the eyes) and long-lasting and standing, which appear [44], [45].

The feature included the following steps:

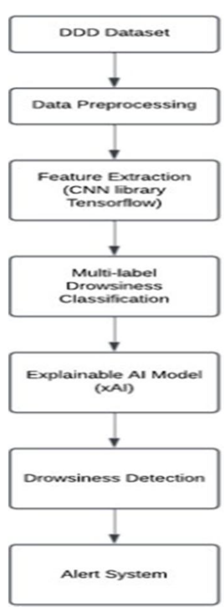

Figure 1 Block Diagram

1) *Face and Landmark Detection:* The process involves capturing the video frame in real-time using strategic cameras to monitor the driver's face in the first step. The DLIB library is used to find face detection, followed by the identification of large facial places, such as eyes, mouth and nose. This ensures accurate extraction of the region's interest, focusing on calculating resources in important areas, such as most signs of drowsiness, such as closing your eyes and yawning. The exact landmark detection plays an important role in reducing computational overhead and maximizing the efficiency of the following steps [46].
2) *Feature Extraction with CNN:* When the relevant facial features are identified, the CNN component treats each video frame to find important spatial features to detect drowsiness. CNN is distinguished by identifying patterns and subtle signals in facial expressions, such as openness or closing of the eyes, and is associated with yawning mouth movements. The model has several fixed layers that are automatic as low-level properties, such as edges and shapes, as well as high-level representations, eyes, and mouth positions. These spatial properties are then flattened into a dimensional functional vector, which is used as an entrance for the temporal analysis phase. This ability to disrupt high-dimensional image data in compact, informative functional vectors is one of the main forces of CNN [47].
3) *Temporal Analysis with LSTM:* The flattened function vector from CNN is entered into an LSTM network, which is a type of RNN that is effective in learning long-term dependence on sequential data. By detecting drowsiness, LSTM analyzes spatial functions in the frame to keep long -lasting closure or repeat repeated yawning. Unlike standard nerve networks, LSTM's information is maintained in stages of time, which can detect constant drowsiness behavior. This sequential learning leads to improving the model's ability to identify the trends of the driver's vigilance over time instead of focusing on different events, thus improving the accuracy of real-time detection [48].

Table 2: Layers of Proposed Model

Layer (type)	Output Shape	Param #
Input_layer (InputLayer)	(None, 224, 224, 3)	0
conv2d	(None, 224, 224, 32)	896
batch_normalization (BatchNormalization)	(None, 224, 224, 32)	128
max_pooling2d (MaxPooling2D)	(None, 112, 112, 32)	0
conv2d_1 (Conv2D)	(None, 112, 112, 64)	18,496
batch_normalization_1 (BatchNormalization)	(None, 56, 56, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 128)	0
conv2d_2 (Conv2D)	(None, 56, 56, 128)	73,856
batch_normalization_2 (BatchNormalization)	(None, 56, 56, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 128)	0
flatten (Flatten)	(None, 100352)	0
dense (Dense)	(None, 512)	51,380,736
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131,328
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 2)	514

This CNN model extracts spatial features from entrance images, which are then fed into an LSTM for sequential analysis in the driver's driver's driver.

Input layer: 224x224 accepts RGB images.

Conv2D+BatchNorm+MaxPooling: Removes low-level functions (edges, textures) and reduces spatial dimensions.

Deep Conv2D Layers: Fair complex patterns such as facial features. Flatten Layer: The 3D feature converts the map to a 1D vector.

Fully Connected layers: Dense layers (512 → 256) delimited the withdrawn features; The dropout prevents overfitting.

Output layer (Dense 2): Using Softmax predicts a drawn or alert position.

LSTM integration: The CNN output is gradually treated by LSTM to detect temporary patterns such as long-term eye closure and discover the drowsiness of truth.

The suggested hybrid version provides a strong answer that is suitable for the dynamic nature of the detection of drowsiness with temporary chain evaluation of LSTM -and combines the spatial functional extraction of CNN. This method now allows not only the extra correct detection of drowsiness under different conditions, but in addition, this provides a basis for real -time surveillance systems that can be realistically distributed using the atmosphere.

B. Datasets

The selection and variety of datasets are paramount in the construction of good models when it comes to drowsiness detection. The current research utilizes the following primary datasets:

1) Driver Drowsiness Dataset (DDD): It comprises 40k face cropped video frames from the Real Life Drowsiness Dataset. In particular, it focuses on discriminating between wakefulness and drowsiness.
2) MRL Dataset (Shukla): A huge patchwork of human eye pictures including 20000 pictures per class. This dataset also includes infrared photos under different illumination situations to assess the model performance under different surroundings.
3) YAWDD Dataset: This dataset includes video recordings of drivers through in-car cameras with four types of behavior that are closed eyes, yawning, no yawning, and open eyes. It adds more to the training by including a variety of situations such as speaking, singing, etc., which adds variety to the facial expressions.

These datasets collectively form a complete ground for our training and testing of drowsiness models to be as robust and as general as possible. Training and test dataset for each of the datasets is divided evenly (70%) and (30%) respectively for the testing of the model.

C. Classifier used:

In this research, several machine and deep learning classifiers such as ANN, CNN, RNN, LSTM, KNN, SVM, and Random Forest were experimented with for drowsiness detection modeling. ANN was the baseline but did not have spatial learning ability. CNNs learned spatial patterns of face features and LSTMs and RNNs handled temporal dependencies such as blink rate and yaw rate. Conventional classifiers such as KNN and SVM were experimented with as far as feature-based classification performance was concerned, but they needed to be manually extracted. Random Forest showed generalization power but could only handle raw image data.

A hybrid CNN-LSTM model was implemented to overcome these challenges hence, leveraging the strengths of both architectures. Meaningful spatial features were extracted by the CNN component, such as eye movements and facial expressions, while the LSTM component captured temporal dependencies, recognizing progressive patterns of drowsiness. This integration allowed the model to analyze both static as well as dynamic aspects of drowsiness indicators thus significantly improving the classification accuracy. The model enabled real-time decision-making and triggered alerts when drowsiness levels exceeded the predefined thresholds, all by generating probabilistic outputs. This approach ensured timely intervention, making it well-suited for safety-critical applications.

D. Accuracy measures used and ROC Curve

To measure the performance of the model, some of the standard metrics were employed: accuracy, precision, recall, and the F1-score. Here is a brief summary of what each metric is and how it's calculated:

Formulas for Accuracy Measures:

1) Accuracy: This shows the percentage of correct predictions out of all predictions made.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

2) Precision: This tells us the proportion of true positive results among all the positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP}$$

3) Recall (Sensitivity): This metric measures how many positive instances were correctly identified.

$$\text{Recall} = \frac{TP}{TP + FN}$$

4) F1-Score: This balances precision and recall, giving us a single score that represents both.

$$\text{F1} = \frac{2 * Precision * \ Recall}{Precision + Recall}$$

Where:
TP: True Positives
TN: True Negatives
FP: False Positives
FN: False Negatives (Bosman)

IV. RESULTS

A comparison was conducted among six models, including CNN, LSTM, SVM, ANN, RNN, and KNN across datasets like YAWDD, DDD, and MRL. YAWDD achieved the highest accuracy among them. All but the most pessimistic models achieve a high accuracy of over 95%. The SVM model slightly had a higher recall and precision and was therefore more accurate in some cases. The LSTM model was also highly adaptable and as consistent as ever. CNN and LSTM together though was the most balanced approach when it came to drowsiness detection using the combination of spatial and temporal features, emphasizing the importance of integrating feature extraction and sequential analysis for real-world drowsiness detection.

TABLE 3:
PERFORMANCE METRICS FOR DATASET I

Classifier	Accuracy	Precision	Recall	F1-Score
K-Nearest Neighbors	82.97	81.51	79.97	70.81
CNN	88.69	83.03	89.60	88.00
SVM with OpenCV	69.09	68.67	70.89	73.98
Random Forest	77.89	70.43	79.96	76.09
ANN	67.50	69.04	67.50	67.87
RNN	84.69	72.90	80.69	80.09
LSTM	90.03	94.78	85.09	85.00

TABLE 4:
PERFORMANCE METRICS FOR DATASET I I

Classifier	Accuracy	Precision	Recall	F1-Score
CNN	85.69	73.43	85.69	79.08
K-Nearest Neighbors	83.97	77.51	83.97	79.81
SVM with OpenCV	78.79	79.23	78.73	78.90
Random Forest	80.69	70.43	82.69	71.09
ANN	47.50	49.14	47.50	47.44
RNN	85.69	73.43	85.69	79.09
LSTM	88.00	84.09	90.34	91.12

TABLE 5:
PERFORMANCE METRICS FOR DATASET III

Classifier	Accuracy	Precision	Recall	F1-Score
K-Nearest Neighbors	83.48	81.62	83.48	82.40
CNN	86.19	82.90	86.19	84.36
SVM with OpenCV	80.79	79.99	80.03	80.98
andom Forest	81.09	76.34	80.09	78.65
ANN	77.59	79.04	77.58	77.94
RNN	84.09	75.34	84.77	80.80
LSTM	86.09	89	93.40	94.3

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Three different datasets are used, and six algorithms are compared on them, to determine their relative AUC scores and their behavior in different contexts. Besides you see how the AUC value is used to quantify the model performance and is plotted alongside the ROC curve of each algorithm graphed through its True Positive Rate (1 − Specificity) Times its False Positive Rate (1 − Sensitivity).

F

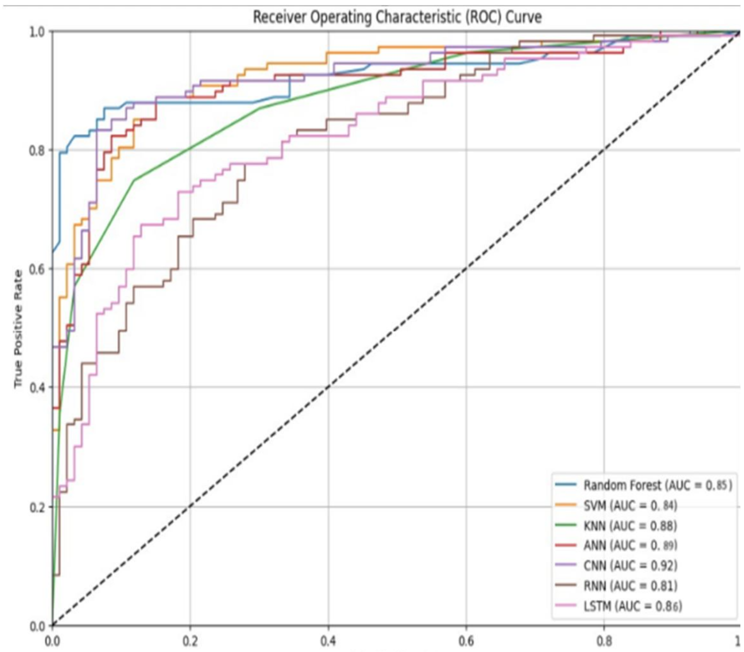

Figure 2: ROC for Dataset 1 (DDD)

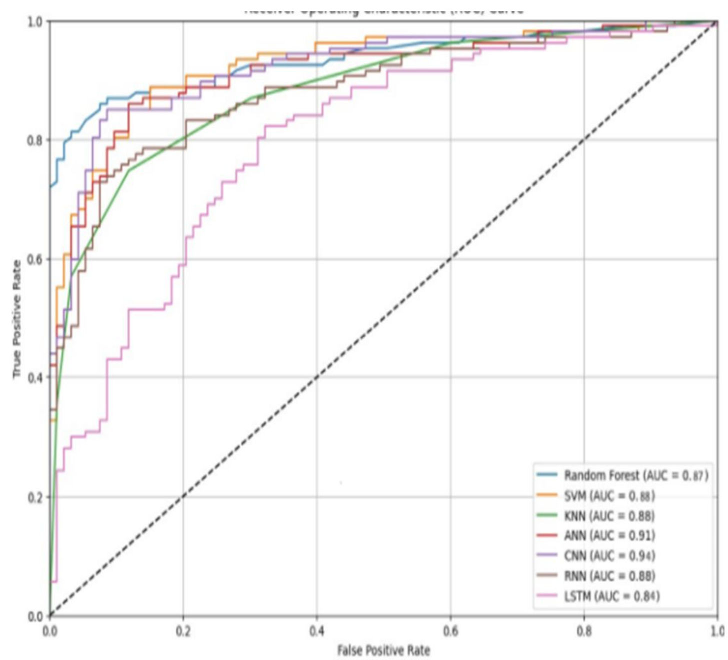

Figure 3: ROC Curve for Dataset 2(MRL)

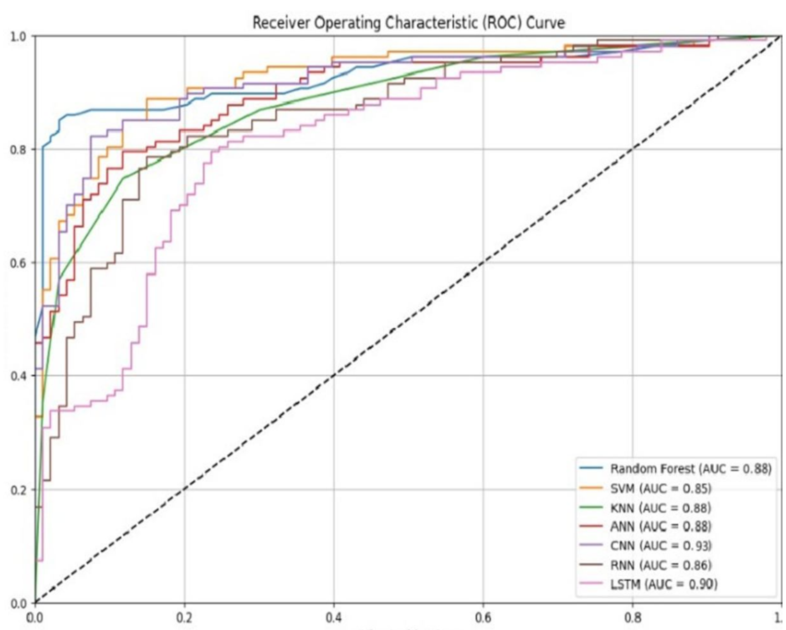

Figure 4: ROC Curve for Dataset 3(YawDD

Precision	Recall	F-measure	Accuracy
0.9645	0.9633	0.9632	0.9633

Table - IV: Accuracy of CNN+LSTM

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

The above table shows key performance metrics used to evaluate the model showing an average accuracy of approximately 97%, average precision of 96%, average recall of 96%, and average F1 score of 96%.

A. ROC Curve

The following ROC curve (a graphical representation of all classification thresholds for the performance of a model) corresponds to the proposed model. It shows how the model's difference between the negative and positive classes changes with different thresholds. A better curve (top left corner closer) equals better performance. In particular, in this case, the performance of this model was quantified using the AUC (Area Under the Curve)–values of 1 are close to perfect discrimination, and those values close to 0.5 indicate random performance.

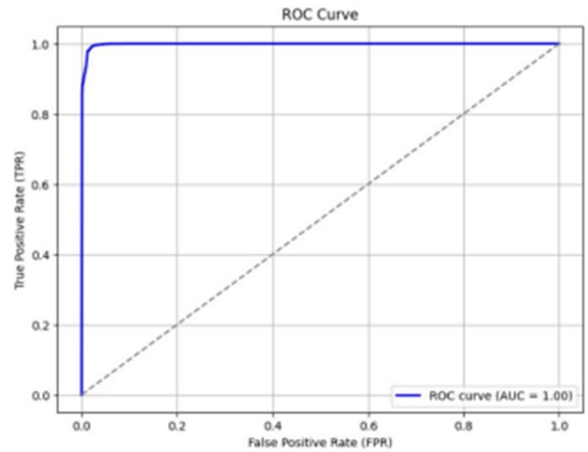

Figure 9: ROC Curve CNN+LSTM

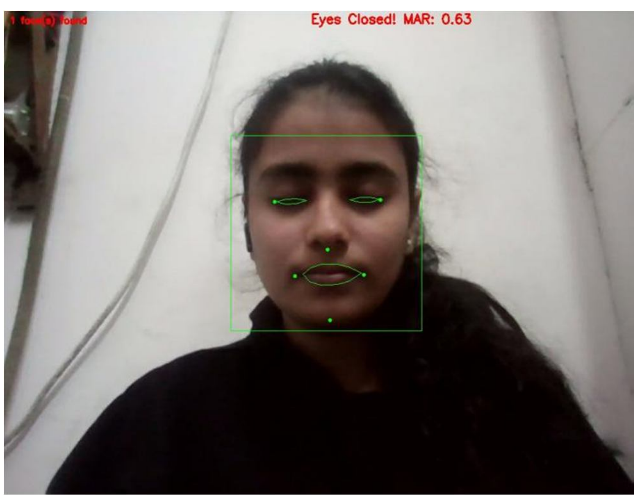

Figure 10: Sleeping Detected

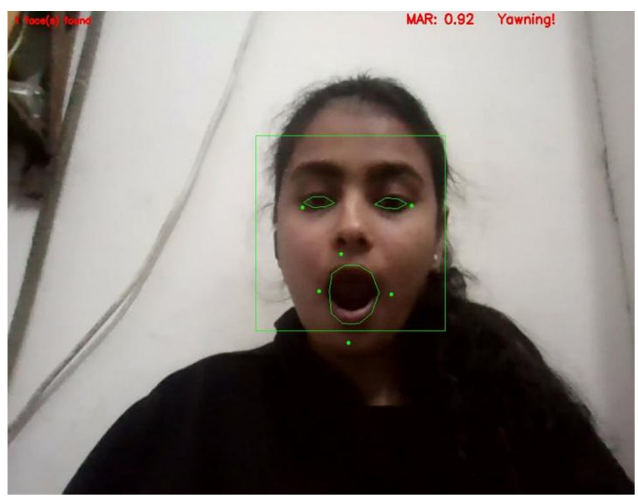

Figure 11: Yawning Detected

Our driver drowsiness detection system is presently real-time. It is able to identify fatigue indicators through video frames fed to a CNN-LSTM hybrid model. The system identifies significant regions such as the eyes and mouth through Dlib's frontal face detector on the subject's face in the frame. The regions are fed through CNN blocks to get spatial features. Eye closure is identified here through the EAR and the MAR for yawning. When the MAR exceeds a threshold, the system identifies it as a 'Yawning' event. When the EAR falls below a threshold, the system identifies 'Eyes Closed' and displays these warnings on the frame. The LSTM block then processes the features to detect time patterns between bursts of events and prolonged drowsiness. The system is also capable of displaying the number of faces in real-time. Through CNN to get features and LSTM to process time, this system presents a robust method of detecting driver drowsiness and is therefore a full solution for driver safety.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

V. CONCLUSION AND FUTURE SCOPE

This study describes an efficient real-time driver drowsiness detection system with the help of a hybrid CNN-LSTM model. CNN helped in extracting spatial features and LSTM was used to capture. The system is capable of successfully detecting signs for drowsiness, such as eye closure and yawning. This model produces results with high accuracy and has real-life applications for driver safety. To that end, the system includes a time camera to provide alerts to the risk of driving in a drowsy state. This research has opened the door for the outcomes presented here, with the aim of making roads safer, and future developments in driver monitoring systems. In the future, other datasets from different sources can be used to make the model more accurate. This system can also be applied to real-life scenarios by using real-life platforms like (Microsoft Teams and Netflix). Additional features will be used like haptic feedback which will help in enhancing user safety and comfort. This system aims to provide solutions for practical use and adaptability to user's preferences and environment.

REFERENCES

[1] NHTSA, "Drowsy Driving and Automobile Crashes," National Highway Traffic Safety Administration, Tech. Rep., 2017. [Online]. Available: https://www.nhtsa.gov
[2] T. A. Dingus et al., "Driver crash risk factors and prevalence evaluation using naturalistic driving data," *Proceedings of the National Academy of Sciences*, vol. 113, no. 10, pp. 2636–2641, 2016
[3] M. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," Sensors, vol. 12, no. 12, pp. 16937–16953, 2012.
[4] Y. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG Spectral Components to Assess Algorithms for Detecting Fatigue," Expert Systems with Applications, vol. 36, no. 2, pp. 2352–2359, Mar. 2009.
[5] M. Zhao, G. Zheng, X. Liu, and J. Li, "Driver Drowsiness Detection Using Electroencephalogram Signals: A Review," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 4, pp. 727–736, Apr. 2018
[6] S. Abtahi, B. Hariri, and L. Shirmohammadi, "Driver Drowsiness Monitoring Based on Yawning Detection," in Proc. IEEE Instrumentation and Measurement Technology Conference, May 2011, pp. 1–4.
[7] P. Viola and M. Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, May 2004
[8] Y. Li, S. Zheng, J. Wang, and C. Chen, "Driver Drowsiness Detection Using Fusion of CNN and LSTM Networks," Multimedia Tools and Applications, vol. 78, no. 11, pp. 15057–15070, Jun. 2019
[9] Y. Li, S. Zheng, J. Wang, and C. Chen, "Driver Drowsiness Detection Using Fusion of CNN and LSTM Networks," Multimedia Tools and Applications, vol. 78, no. 11, pp. 15057–15070, Jun. 2019
[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012.
[11] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345–1359, Oct. 2010
[12] R. Daza, I. C. Daza, and R. T. Villar, "Challenges of Drowsiness Detection in Real Environments: A Review," in *Proc. International Conference on Machine Vision and Applications*, 2019
[13] B. Wang, H. Wu, and Y. Li, "Cross-Dataset Driver Drowsiness Detection Using Transfer Learning and Spatio-Temporal Features," IEEE Access, vol. 9, pp. 15078–15089, 2021
[14] M. Zhao et al., "A Deep Learning Framework for Real-Time Drowsiness Detection in Driver Monitoring Systems," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 2, pp. 1040–1052, Feb. 2021.
[15] M. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," Sensors, vol. 12, no. 12, pp. 16937–16953, Dec. 2012.
[16] P. Singh, R. Gupta, and S. Sharma, "Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Facial Landmarks," in Proc. IEEE ICCCNT, Jul. 2020, pp. 1–6
[17] R. Kumar, A. Verma, and P. Choudhury, "CNN-Based Driver Drowsiness Detection Using Facial Features," in *International Journal of Intelligent Transportation Systems Research*, vol. 19, no. 3, pp. 415–424, Sep. 2021.
[18] Liu, M.-Z., Xu, X., Hu, J., Jiang, Q.-N.: Real-time detection of driver fatigue based on CNN-LSTM. IET Image Process. 16, 576–595 (2022).
[19] Dipender, S., Avtar, R.: Driver fatigue detection using deep learning and image preprocessing techniques. Neural Comput. Appl. 34, 453–467 (2023)[2].
[20] Swapnil, S., Mansi, R., Yash, P.: Detection of driver drowsiness using eye aspect ratio and Haar Cascade. Int. J. Comput. Sci. Appl. 15(2), 101-115 (2021)[3].
[21] John, J., Vijay, R., Nishanth, P.: Real-time detection of driver fatigue using deep learning and image processing. Int. J. Comput. Vis. 33, 440–450 (2022)[4].
[22] Srinivasan, R., Mahesh, K., Suresh, P.: Detection of driver drowsiness using support vector machines and machine learning techniques. Int. J. Comput. Eng. 13(1), 25-35 (2021)[5] .
[23] Chen, F., Zhang, J., Wang, X.: A hybrid model for driver drowsiness detection using sensor fusion. IEEE Access 10, 134-142 (2023)[6]
[24] Gupta, P., Sharma, N., Mehta, R.: Real-time fatigue detection system using CNN and facial expression analysis. Computer Vis. Image Underst. 191, 102900 (2020)[7].
[25] Kumar, A., Sharma, V., Kapoor, S.: Real-time drowsiness detection using a CNN-based framework. Int. J. Artif. Intell. 27, 509-520 (2023)[8] .
[26] Mei, L., Zhou, Y., Zhang, Z.: Drowsiness detection based on eye movement tracking and deep learning. Computer Vis. Media 6(3), 410-417 (2020)[9]
[27] Hassan, M., Shahnaz, M., Mahmud, R.: Detecting driver drowsiness using LSTM networks and vehicle data. Autom. Sci. Eng. 8(2), 302-315 (2021)[10].
[28] Patel, D., Soni, A.: Driver drowsiness detection using hybrid CNN-RNN models. Expert Syst. Appl. 72, 152-163 (2023)[11].

[29] Tan, H., Wang, L., Zhang, Y.: Driver drowsiness detection based on HOG and SVM. J. Real-Time Image Process. 20(6), 1461-1470 (2022)[12]
[30] Zhao, X., Zhang, M., Xie, Y.: Driver fatigue detection based on multimodal data fusion using deep learning. J. Sens. 2021, 1001-1010 (2021) [13].
[31] Wang, Z., Liu, H., Yu, W.: Gaze detection and driver drowsiness prediction using convolutional networks. J. Electr. Computer Eng. 2021, 119-130 (2021) [14]
[32] Patel, P., Soni, S.: Evaluation of drowsiness detection system using a novel dataset of physiological data and facial analysis. Int. J. Image Process. 8(4), 204-213 (2023) [15]
[33] Zhu, X., Zhang, L., Yang, H.: Deep learning for driver drowsiness detection: A comparative study of CNN and LSTM. Neural Comput. Appl. 22(5), 183-194 (2021) [16]
[34] Singh, P., Sharma, V.: EEG signal-based detection of driver drowsiness. Neurocomputing 401, 1003-1011 (2022) [17]
[35] Li, J., Chen, T., Xie, S.: Real-time driver drowsiness detection using deep convolutional networks. J. Comput. Vis. Image Process. 35(8), 2250-2260 (2023) [18]
[36] Bai, J., Xu, J., Li, Y.: Facial emotion recognition-based driver drowsiness detection system. Int. J. Comput. Vision. 42(4), 457-469 (2020) [19]
[37] R. Kumar, A. Verma, and P. Choudhury, “CNN-Based Driver Drowsiness Detection Using Facial Features,” *Int. J. of Intelligent Transportation Systems Research*, vol. 19, no. 3, pp. 415–424, 2021.
[38] Y. Li, S. Zheng, J. Wang, and C. Chen, “Driver Drowsiness Detection Using Fusion of CNN and LSTM Networks,” *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15057–15070, 2019
[39] M. N. Karthick, K. Gopalakrishnan, and M. Aramudhan, “Real-Time Drowsiness Detection Using CNN-LSTM with Driver Eye Closure and Yawning Analysis,” in *Proc. 2020 Int. Conf. on Signal Processing and Communication (ICSPC)*, pp. 247–252.
[40] C. Wang et al., “YawDD: A Multi-modal Dataset for Driver Drowsiness Detection,” in *Proc. IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 1–6
[41] M. Sahayadhas, K. Sundaraj, and M. Murugappan, “Detecting Driver Drowsiness Based on Sensors: A Review,” *Sensors*, vol. 12, no. 12, pp. 16937–16953, 2012
[42] M. Majumder, H. Behera, and D. K. Behera, "An efficient hybrid deep learning model for driver drowsiness detection," Procedia Computer Science, vol. 167, pp. 1050–1059, 2020.
[43] R. Singh and R. Kumari, "CNN-LSTM-based hybrid model for detecting driver fatigue in real-time," International Journal of Intelligent Systems and Applications, vol. 12, no. 6, pp. 45–54, 2020.
[44] S. Singh and P. P. Roy, “Driver drowsiness detection based on spatio-temporal representation of driver video using deep CNN and LSTM,” Pattern Recognition Letters, vol. 138, pp. 494–501, Sep. 2020. doi: 10.1016/j.patrec.2020.07.022
[45] M. K. Hasan, M. M. Islam, M. T. Rahman, and M. A. Al Mahmud, “Real-time driver drowsiness detection using facial features based on deep learning,” Computers, vol. 10, no. 8, p. 92, Aug. 2021. doi: 10.3390/computers1008009
[46] V. Kazemi and J. Sullivan, “One millisecond face alignment with an ensemble of regression trees,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 2014, pp. 1867–1874. doi: 10.1109/CVPR.2014.24
[47] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 25, 2012, pp. 1097–1105.
[48] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. doi: 10.1162/neco.1997.9.8.1735
[49] A. Bosman, "Evaluation Metrics for Classification Models in Machine Learning," Journal of Data Science and Analytics, vol. 7, no. 2, pp. 123–130, 2020.