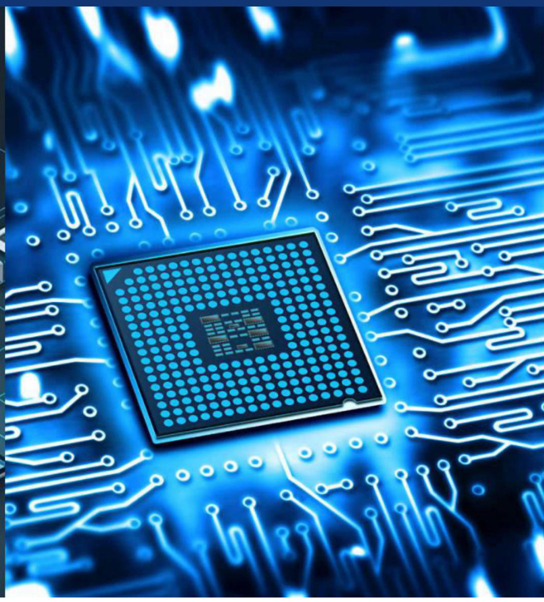
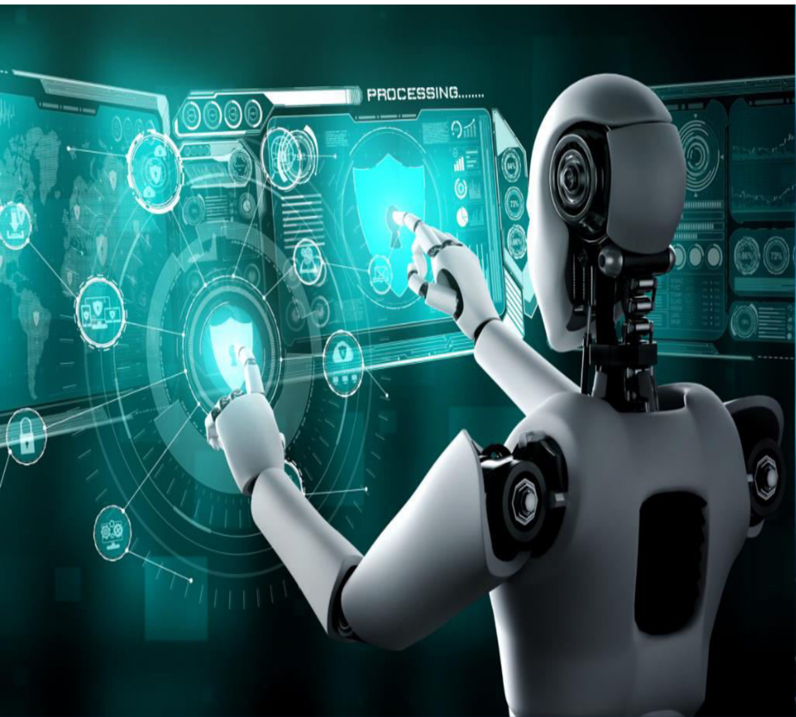
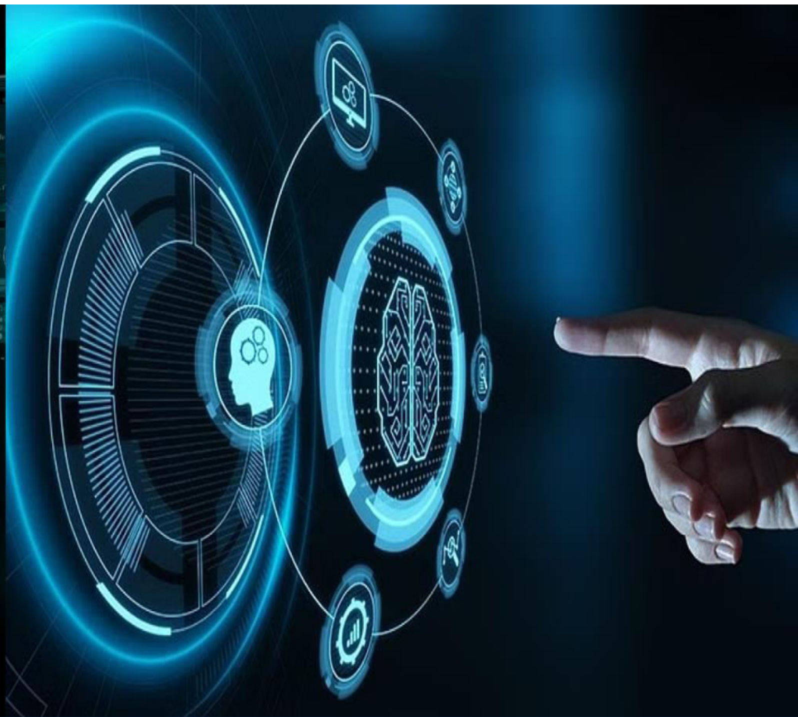
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 5, May 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Design and Implementation of Driver Drowsiness Detection System using Open CV

K. Nandini, M. Bhavya, M. Keerthana, J. Sowmiya

Assistant Professor, Department of CSE, P.S.V College of Engineering and Technology, Mittapalli, Krishnagiri, India

UG Scholar, Department of CSE, P.S.V College of Engineering and Technology, Mittapalli, Krishnagiri, India

UG Scholar, Department of CSE, P.S.V College of Engineering and Technology, Mittapalli, Krishnagiri, India

UG Scholar, Department of CSE, P.S.V College of Engineering and Technology, Mittapalli, Krishnagiri, India

ABSTRACT - Road accidents due to driver fatigue and drowsiness have become a significant concern worldwide, leading to loss of lives, property damage, and economic burdens. A Driver Drowsiness Detection System (DDS) is an advanced safety mechanism designed to prevent accidents by monitoring a driver's alertness and providing timely warnings. This paper presents an efficient and real-time drowsiness detection system that integrates machine learning, image processing, and sensor-based methodologies to enhance road safety. The proposed system employs a combination of facial recognition, eye state monitoring, and head movement tracking to determine signs of fatigue. Using a camera positioned in front of the driver, real-time video frames are captured and analyzed to detect eye closure duration, yawning frequency, and head tilting, which are indicative of drowsiness. Advanced image processing techniques, such as Haar cascades and deep learning-based Convolutional Neural Networks (CNNs), are utilized for accurate face and eye feature extraction. Additionally, machine learning models are trained on datasets containing various drowsiness indicators to improve the system's accuracy. To complement image-based detection, physiological sensors, such as heart rate and skin conductance sensors, are integrated to capture biometric signals that correlate with drowsiness. The system continuously processes these multimodal data sources and applies fusion techniques to enhance prediction reliability. When the system detects signs of drowsiness, it issues real-time alerts through auditory alarms, visual notifications, or haptic feedback to ensure the driver regains focus. The effectiveness of the proposed DDS is evaluated using benchmark datasets, such as the NTHU Drowsy Driver Dataset, along with real-world testing in simulated and on-road environments. Experimental results demonstrate high accuracy in detecting drowsiness, surpassing conventional methods that rely solely on eye blink frequency. The system is designed for integration with in-vehicle infotainment systems, making it suitable for commercial and personal use.

KEYWORDS: Men Driver Drowsiness Detection, Machine Learning, Image Processing, Convolutional Neural Networks, Sensor Fusion, Road Safety, Artificial Intelligence.

I. INTRODUCTION

Drowsy driving is a leading cause of road accidents, resulting in fatalities and financial losses. Existing detection systems mainly focus on eye closure, ignoring other signs like yawning. This project proposes an enhanced detection system using CNN and facial landmark analysis. The system integrates both eye closure and yawning detection for improved accuracy. A real-time buzzer alert system is included to wake the driver instantly. The proposed model ensures higher accuracy and responsiveness compared to traditional methods. Designed for real-world deployment, it enhances driver safety and reduces accident risks.

II. METHODOLOGY

The proposed Drowsy Detect Net framework is designed to detect driver drowsiness efficiently using a lightweight and shallow CNN model that performs well with limited training data. The methodology includes the following key steps:

1. Face and Eye Region Detection:

- A 68-point facial landmark detection algorithm (from the Dlib library) is used to locate facial features.

ii Only the eye regions are considered critical for drowsiness detection, particularly eyelid closure.
iii Eye regions are extracted by identifying landmarks:
iv Left eye: Points 37–42
v Right eye: Points 43–48
vi These regions are cropped and used as the input for further analysis.

2. Shallow CNN Architecture:
A custom shallow CNN was developed for binary classification (open vs. closed eyes).
Architecture Details:
i Convolutional Layers
a. Filters: 32, 64, 128 (3×3), and 128 (1×1)
b. Each followed by a 2×2 Max-Pooling layer
c. Dropout Layers (Dropout rate: 0.2)
ii Fully Connected Layers:
a. 128 nodes each
b. Final layer uses Sigmoid activation for binary output
Activation Functions:
a.ReLU for intermediate layers
b.Sigmoid for output
c.Input image size for eye regions: 128 × 128 pixels

3. Dataset Utilization:
a. Dataset-1: 324 images (custom collected)
b. Dataset-2: 1452 images from Kaggle Yawn_Eye_Dataset_New
c. Both datasets consist of labeled open and closed eye images
d. Data was split into training (60%) and testing (40%)

4. Hyperparameter Tuning:
a. Learning Rate: 0.001 yielded the best results
b. Epochs: 100 epochs provided optimal performance
c. Batch Size: 32 performed better in terms of speed and accuracy
d. Optimizer: Adam performed better than alternatives like SGD and RMSprop

5. Performance Metrics:
Evaluated using:
Accuracy
a. Precision
b. Recall
c. F1-Score
d. ROC and PR Curves
Achieved high accuracy:
a.99.23%* on Dataset-1
b.99.14%* on Dataset-2
Confusion matrices confirm the robustness of classification in both open and closed eye categories

6. Comparison with Pre-Trained Models:
The proposed model outperformed heavier architectures such as:
i VGG19,ResNet50,InceptionV3,and MobileNetV2
It had:
a. Lower inference time
b. Fewer parameters
c. Greater suitability for real-time embedded systems

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. DISADVANTAGE OF EXISTING SYSTEM

1. Many existing systems struggle to accurately interpret subtle expressions, idiomatic phrases, or complex emotional cues, leading to misinterpretation or irrelevant responses.
2. Reliance on predefined templates often results in generic, repetitive answers that lack personalization and fail to address individual user needs effectively.
3. When faced with intricate or multi-faceted medical questions, existing systems may falter, leading to incomplete, shallow, or inaccurate guidance.
4. As the number and complexity of user interactions grow, many systems experience performance bottlenecks, resulting in slower responses, reduced accuracy, and diminished user satisfaction.

III. DESIGN AND ARCHITECTURE

The proposed Drowsy Detect Net framework is a lightweight and efficient architecture tailored to detect driver drowsiness using visual features, primarily focused on the eye region. This section outlines the overall system design and the detailed architecture of the shallow CNN model used for binary classification of eye states.

A. System Design Overview:

1. Image Acquisition:
   A camera mounted on the dashboard continuously captures video frames of the driver's face in real time.
2. Facial Landmark Detection:
   Each frame is passed through a facial landmark detection algorithm based on the Dlib library, which extracts **68 facial key points**. Key indices corresponding to the eyes (points 37–42 for the left eye and 43–48 for the right eye) are used to locate the eye region.
3. Region of Interest (ROI) Extraction:
   The eye regions are cropped using the identified coordinates and resized to **128×128 pixels**. This normalization facilitates consistent input into the CNN classifier.
4. Drowsiness Detection using CNN:
   The eye images are passed into a custom-designed shallow Convolutional Neural Network. The CNN processes the features and classifies each eye as either "open" or "closed."
5. Decision Module:
   A decision mechanism monitors the eye states over successive frames. If both eyes remain closed beyond a predefined time threshold (e.g., 3 seconds), the driver is flagged as drowsy, and an alert is triggered.

B. Proposed CNN Architecture

The Drowsy Detect Net uses a shallow CNN with reduced depth and parameters, suitable for systems with limited computing resources. The model's lightweight structure makes it ideal for real-time deployment in embedded platforms.

Architectural Components:

Layer	Details
Input Layer	128×128 grayscale image (cropped eye region)
Conv2D-1	32 filters, 3×3 kernel, ReLU activation
MaxPooling2D-1	2×2 pool size
Dropout-1	0.2 dropout rate
Conv2D-2	64 filters, 3×3 kernel, ReLU activation
MaxPooling2D-2	2×2 pool size
Dropout-2	0.2 dropout rate
Conv2D-3	128 filters, 3×3 kernel, ReLU activation
MaxPooling2D-3	2×2 pool size
Conv2D-4	128 filters, 1×1 kernel, ReLU activation
MaxPooling2D-4	2×2 pool size
Flatten	Converts feature maps to a 1D feature vector
Dense-1	128 nodes, ReLU activation
Output (Dense-2)	1 node, Sigmoid activation (binary classification)

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

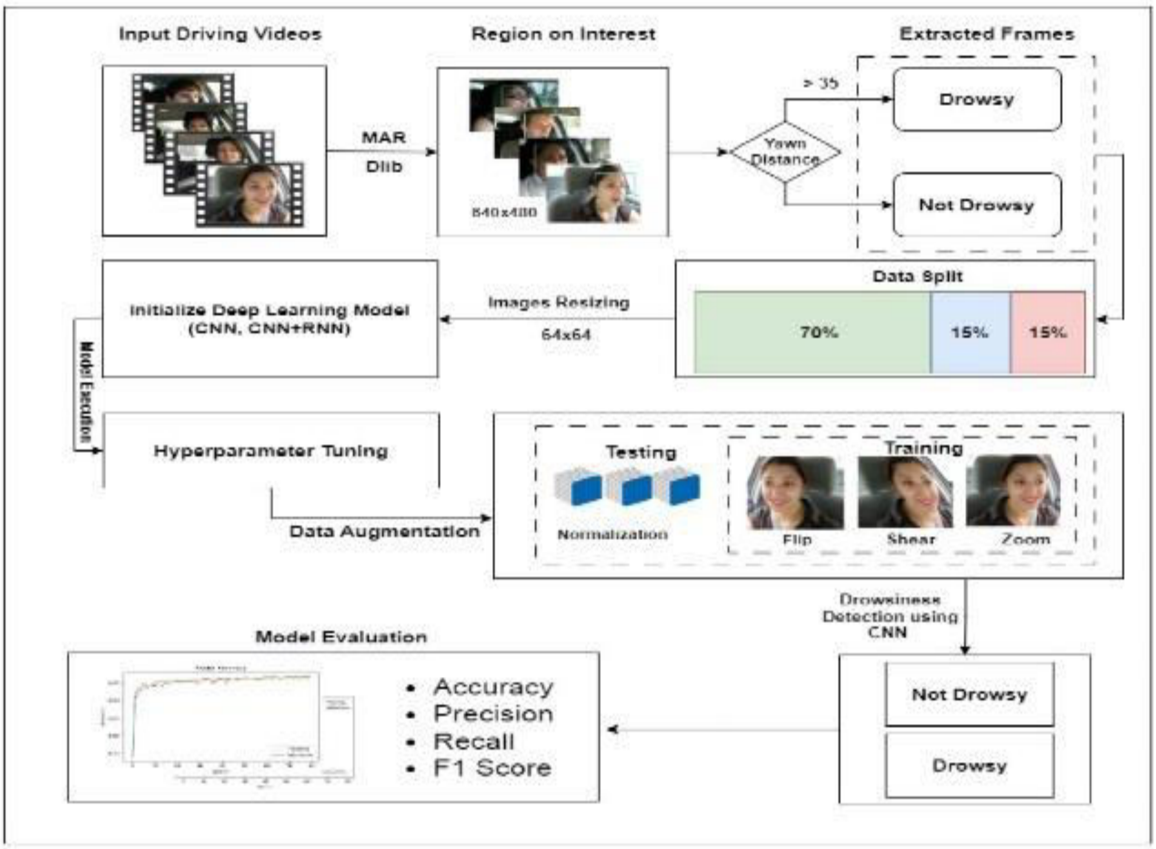
Activation Functions:
I ReLUis used in all convolutional and dense layers except the output.
II Sigmoidis used in the output layer to classify the eye state as open or closed.

C. System Flowchart:
The logical flow of the Drowsy Detect Net framework is illustrated below:

Input Video Frame
↓
Facial Landmark Detection (Dlib 68 Points)
↓
Eye Region Extraction (Points 37–48)
↓
Preprocessing (Resize to 128×128)
↓
Shallow CNN Inference (Open/Closed)
↓
Temporal Decision Logic (Track over Time)
↓
Drowsiness Alert (If Eyes Closed > Threshold)

This design enables high-accuracy eye-state classification with significantly reduced computational complexity, making it suitable for integration in real-time automotive driver monitoring systems.

Fig 1: Architecture Diagram



IV. IMPLEMENTATION / EXPERIMENTAL SETUP

This section outlines the implementation details, hardware/software environment, dataset preparation, training configuration, and evaluation metrics used to assess the performance of the proposed DrowsyDetectNet framework.

A. Hardware and Software Environment:

1. Hardware:

a. CPU: Intel Core i5 Processor (9th Gen)
b. RAM: 16 GB DDR4
c. GPU: NVIDIA GeForce GTX 1050Ti (4 GB)

2. Software:

a. Operating System: Windows 10 (64-bit)
b. Programming Language: Python 3.9

3. Libraries and Frameworks:

a. TensorFlow 2.x
b. Keras
c. OpenCV
d. Dlib (for 68-point facial landmark detection)
e. NumPy, Matplotlib, scikit-learn

B. Dataset Description:

Two datasets were utilized for training and evaluation:

1.Dataset-1:
a. Custom dataset created by the authors.
b. Total images: 324 (balanced between open and closed eyes).
c. Format: RGB, manually labeled.

2. Dataset-2:
d. Public dataset: Yawn_Eye_Dataset_New from Kaggle.
e. Total images: 1452 (balanced classes).
f. Image categories: “open eyes” and “closed eyes”.

3.Data Splitting:
a. Both datasets were split into **training (60%)** and **testing (40%)** sets.
b. During training, 20% of the training data was further used for validation.

C. Preprocessing and ROI Extraction:

i.Each image was processed using the Dlib library to detect 68 facial landmarks.
ii.Eye regions were identified using indices:
iii.Left Eye: 37–42
iv.Right Eye: 43–48
v.The identified eye regions were cropped and resized to 128 × 128 pixels to standardize the CNN input.
vi.All images were converted to grayscale for reduced computational cost.

D. CNN Training Configuration:

The proposed shallow CNN architecture was trained with the following hyperparameters:

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Epochs	100
Batch Size	32
Dropout Rate	0.2
Loss Function	Binary Cross-Entropy
Activation	ReLU (hidden), Sigmoid (output)

E. Evaluation Metrics:

To assess the performance of the model, the following metrics were used:

a. Accuracy: Overall percentage of correct predictions.
b. Precision: Correct positive predictions over all positive predictions.
c. Recall: Correct positive predictions over all actual positives.
d. F1-Score: Harmonic mean of precision and recall.
e. ROC Curve and AUC: Performance across different threshold values.
f. Confusion Matrix: Breakdown of TP, TN, FP, FN.

F. Experimental Results Summary:

The proposed DrowsyDetectNet achieved the following:

1. Dataset-1:

a. Accuracy: 99.23%
b. F1-Score: 0.99

2. Dataset-2:

c. Accuracy: 99.14%
d. F1-Score: 0.99

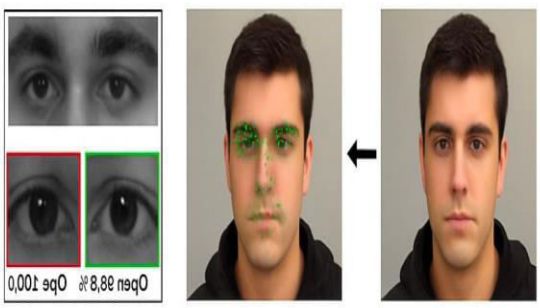
These results outperformed deep learning baselines such as VGG19, ResNet50, MobileNetV2, and InceptionV3 in both accuracy and inference time.

V. RESULT AND DISCUSSION

Eye Detection and ROI Extraction:

The proposed DrowsyDetectNet utilizes Dlib's 68-point facial landmark detection to locate and crop the eye region from input facial images. Figure 1 illustrates this step-by-step process.

Fig 2: Eye Detection and ROI Extraction Process



Figur 1 Eye Detection and Cropping Process

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Eye State Classification Output:

After extracting eye regions, the shallow CNN model classifies them as either "Open" or "Closed." Figure 2 shows a grid of test images with predicted eye states and associated confidence levels.

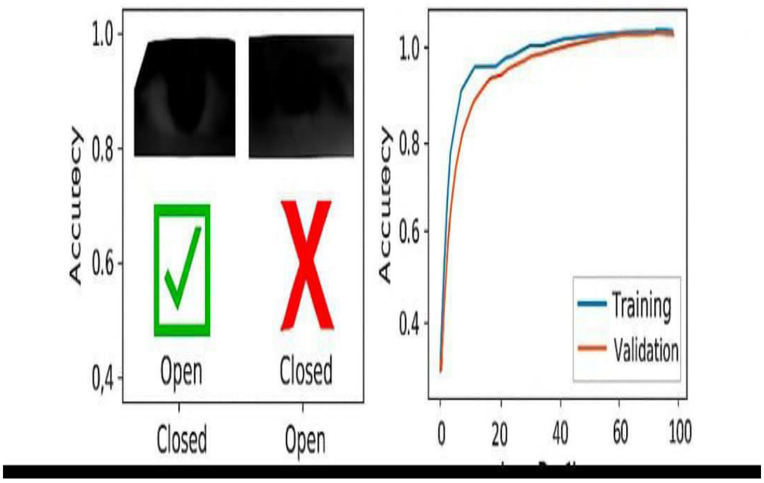
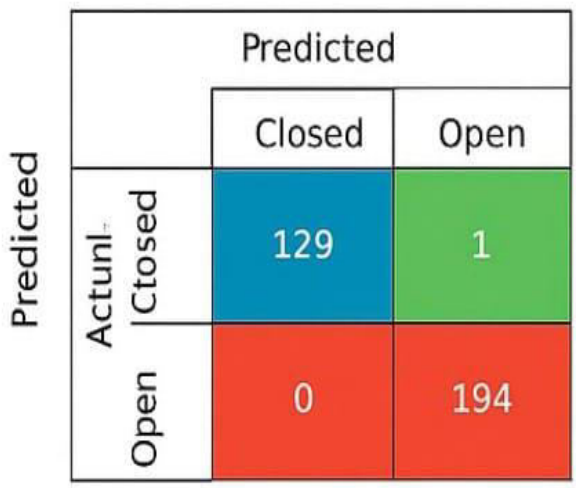


Fig 3: Eye State Prediction Output

Performance Evaluation:

Confusion Matrix: Figure 3 displays the confusion matrix for Dataset-1. The matrix shows near-perfect classification, confirming the model's capability in binary classification tasks (Open vs Closed eyes).



Figur 3 Confusion Matrix
for Dataset-1

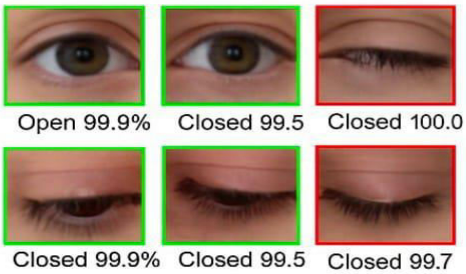
Fig 4: Confusion Matrix on Dataset-1

(High TP and TN, minimal FP and FN) Accuracy Graph: Figure 4 shows the training and validation accuracy over 100 epochs. The graph indicates smooth convergence and generalization.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Figur 2 Eye State Classification Res-

Fig 5: Training vs Validation Accuracy Curve

Quantitative Results:

Metric	**Dataset-1**	**Dataset-2**
Accuracy	99.23%	99.14%
Precision	99%	99%
Recall	100%	99%
F1-Score	0.99	0.99
Inference Time	1430 ms	164 ms

Comparative Analysis:
Compared with standard deep models, the proposed shallow CNN performs better in both accuracy and speed:

Model	Accuracy (D1)	Accuracy (D2)
VGG19	98.46%	98.62%
ResNet50	98.46%	98.10%
InceptionV3	96.15%	93.12%
MobileNetV2	97.69%	91.06%
Shallow CNN	99.23%	99.14%

Discussion:
The proposed system demonstrates that a shallow CNN architecture can yield exceptional accuracy for real-time drowsiness detection, especially with limited training data. Its minimal complexity ensures low latency, making it ideal for integration into vehicle-based monitoring systems.
However, current limitations include:
a. Dataset diversity (e.g., different lighting, demographics)
b. Limited facial cues (focuses only on eye state)

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. CONCLUSION

In this study, a lightweight and efficient driver drowsiness detection framework named DrowsyDetectNet was proposed. The system leverages a shallow Convolutional Neural Network (CNN) architecture that focuses on critical visual cues, particularly eyelid closure, to determine the driver's alertness level. The use of a 68-point facial landmark detector enabled precise eye region extraction, enhancing classification accuracy while maintaining low computational overhead.

Comprehensive experiments conducted on two datasets—one custom and one public—demonstrated the robustness of the proposed model. The shallow CNN achieved 99.23% accuracy on Dataset-1 and 99.14% on Dataset-2, outperforming several deep learning models including VGG19, ResNet50, InceptionV3, and MobileNetV2. Additionally, the model exhibited significantly faster inference times, making it suitable for real-time deployment in automotive environments.

Despite its high performance, the system has some limitations, such as reduced robustness under poor lighting or in the presence of occlusions (e.g., sunglasses). To address these challenges, future work will focus on:Expanding the dataset to include varied lighting conditions and facial features, Incorporating additional behavioral cues such as yawning and head pose, Exploring multimodal approaches including physiological data, Adapting the model for low-power embedded platforms. Overall, DrowsyDetectNet presents a promising solution for enhancing road safety through real-time, accurate, and resource-efficient drowsiness detection.

VII. FUTURE WORK

While the proposed DrowsyDetectNet framework has demonstrated high accuracy and efficiency in detecting driver drowsiness using a shallow CNN model, several areas remain for future exploration and enhancement:

Dataset Expansion and Diversity:
The current datasets primarily contain well-lit, front-facing images with limited demographic and environmental variety. Future efforts will focus on collecting or integrating more comprehensive datasets that include:
a. Night-time and low-light conditions
b. Drivers of various ethnicities and age groups
c. Variations in head pose and facial occlusion (e.g., sunglasses, masks)

Multimodal Drowsiness Detection:
Eye state alone, while effective, may not be sufficient in all scenarios. Combining additional behavioral and physiological indicators can significantly improve robustness:
a. Yawning detection
b. Head pose estimation and nodding behavior
c. Heart rate, EEG, or other biosignal integration using wearable sensors

Real-World Testing and Deployment:
The current system was evaluated in a controlled environment. Future work includes:
a. Deploying the model on embedded platforms such as Raspberry Pi or NVIDIA Jetson Nano
b. Testing in real driving conditions across varied terrains and traffic situations
c. Integration with vehicle infotainment or safety alert systems

Personalized Drowsiness Models:
Driver fatigue patterns vary between individuals. Incorporating adaptive or personalized learning mechanisms could help tailor the model to individual driver behavior over time.
Robustness to Occlusion and Distraction:
a. Enhancing model reliability in cases of partial facial visibility—such as when the driver is wearing sunglasses or looking sideways—will be addressed using advanced techniques such as:
b. Data augmentation for occlusion
c. Incorporating temporal sequence modeling using LSTM or Transformers

By addressing these areas, the DrowsyDetectNet system can be evolved into a fully reliable, real-world solution for improving road safety through proactive fatigue monitoring.

REFERENCES

1. P.Viola & Jones.M. “Rapid object detection using a boosted cascade of simple features” https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf,2020
2. Soukupova.T & Cech.J. “Real time eye blink detection using facial landmark”https://vision.fe.unilj.si/cvww2016/proceedings/papers/05.pdf,2021
3. Abtahi.Set.al. “Yawning detection using embedded smart cameras” https://doi.org/10.1109/ICME.2014.6890255,2023
4. W. W. Wierwille,L. A. Ellsworth.“Developed to Combines yawning and eye closure features to build a more robust system” https://doi.org/10.1016/0001-4575(94)90040-X,2023
5. Md.Nazmul Hasan, Farhana Haque. “Real time drowsiness detection using facial landmark” https://doi.org/10.5120/ijca2019918804,2024
6. M. Dua, R. Singla, S. Raj, and A. Jangra, ‘‘Deep CNN models-based ensemble approach to driver drowsiness detection,’’ Neural Comput. Appl., vol. 33, no. 8, pp. 3155–3168, Apr. 2021.
7. J. Flores-Monroy, M. Nakano-Miyatake, G. Sanchez-Perez, and H. PerezMeana, ‘‘Visual-based real time driver drowsiness detection system using CNN,’’ in Proc. 18th Int. Conf. Electr. Eng., Comput. Sci. Autom. Control (CCE), Nov. 2021, pp. 1–5.., vol. 33, no. 8, pp. 3155–3168, Apr. 2021.
8. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, ‘‘MobileNetV2: Inverted residuals and linear bottlenecks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 4510–4520.
9. V. Chirra, S. ReddyUyyala, and V. Kishorekolli, ‘‘Deep CNN: A machine learning approach for driver drowsiness detection based on eye state,’’ Revue d’Intell. Artificielle, vol. 33, no. 6, pp. 461–466, Dec. 2019.
10. B. Ganguly, D. Dey, and S. Munshi, ‘‘An integrated system for drivers’ drowsiness detection using deep learning frameworks,’’ in Proc. IEEE VLSI Device Circuit Syst. (VLSI DCS), Feb. 2022, pp. 55–59.
11. Z.-T. Liu, C.-S. Jiang, S.-H. Li, M. Wu, W.-H. Cao, and M. Hao, ‘‘Eye state detection based on weight binarization convolution neural network and transfer learning,’’ Appl. Soft Comput., vol. 109, Sep. 2021, Art. no. 107565
12. M. A. R. Ahad, S. Kobashi, and J. M. R. Tavares, ‘‘Advancements of image processing and vision in healthcare,’’ J. Healthcare Eng., vol. 2018, Jan. 2018, Art. no. 8458024.
13. M. Papakostas, K. Das, M. Abouelenien, R. Mihalcea, and M. Burzo, ‘‘Distracted and drowsy driving modeling using deep physiological representations and multitask learning,’’ Appl. Sci., vol. 11, no. 1, p. 88, Dec. 2020.
14. D. Li, Z. Wang, Q. Gao, Y. Song, X. Yu, and C. Wang, ‘‘Facial expression recognition based on electroencephalogram and facial landmark localization,’’ Technol. Health Care, vol. 27, no. 4, pp. 373–387, Jul. 2019.