DROWSY DRIVER DETECTION USING HEAD POSTURE: A DEEP LEARNING APPROACH

Nurudeen A. Iromini[1], Abubakar S. Nafiu[2], Adedoyin S, Adeyemi[3], Bilikis T. Ibrahim[4], Munir A. Ibrahim[1], and Suzie M. Isaac[3]

[1] Department of Computer Engineering, Federal Polytechnic Offa,
[2] Department of Renewable Energy, Federal Polytechnic Offa
[3] Department of Artificial Intelligence, Federal Polytechnic Offa
[4] Department of Software and Web Development, Federal Polytechnic Offa

Email: ***ajibolaforyu2003@yahoo.com, +234 806 298 5344***

Abstract

Road accidents due to driver drowsiness are one of the major issues, leading to many deaths and injuries globally. Drowsiness affects a driver's alertness, reaction time, and decision-making, which often resulting into serious accidents. Current methods for detecting drowsiness, like self-reporting or physiological sensors, are not practical or reliable for everyday use. There is a need for a non-intrusive, real-time monitoring system to identify drowsiness before accidents happen. A new system was introduced in this study which uses DLIB's facial landmark predictor for head posture and a ResNet-50 classifier to identify drowsy drivers accurately by assessing facial features like head tilt and eye closure, proving effective in different lighting and angles. Experimental result showed accuracy of 90%, precision of 0.88, recall of 0.88 and Fl-score of 0.88.

Keywords**:** resnet50, drowsy, Cnn, accidents, driver, Opencv

Introduction

The crucial element in the operation and the development of every civilization is transport. A robust and sophisticated transportation and communication infrastructure is required for prioritizing economic growth strategy and fast industrialization. In recent years, the need for contemporary transport has surged because to population expansion (Ghimire et al., 2025). Currently, the essential mode of transporting people and goods from one place to another is automobile. In 2024, global income of light automobiles reached 84 million, with the anticipation of reaching 85 million for 2025 (WHO, 2024). The worldwide automobile industry is valued at over \$3.6 trillion, with projections indicating it will quadruple to almost \$7 trillion by 2033(WHO, 2024). Over 93.9 million new vehicles were manufactured in 2023, whereas about 72 million registered passengers' cars were operational during the year (WHO, 2024). Approximately 1.47 billion passenger vehicles are in use worldwide. Sales of electric vehicles (EVs) and plug-in hybrid electric vehicles (PHEVs) reached around 17.0 million in 2024, constituting 20–25% of worldwide new automobile sales (European Commission, 2024). By mid-2025, worldwide electric vehicle sales surpassed 1.6 million units in May alone, reflecting a 24% year-over-year increase, mostly propelled by China (WHO, 2024).

The WHO estimated that 1.19 million people die each year in road traffic accidents, while 20 to 50 million get non-fatal injuries. More than 90% of these deaths transpire in low- and middle-income nations, despite their possession of just around 60% of the global automobile fleet (Statista, 2024). Road traffic injuries are the primary cause of mortality for those aged 5 to 29 years, with two-thirds of deaths occurring in the 18 to 59 age group. At-risk user pedestrians, cyclists, motorcyclists, and e-scooter operators constitute more than 50% of worldwide road fatalities (Statista, 2024). The economic repercussions of accidents are substantial, estimated at 3% of worldwide GDP, highlighting significant social expenses. In 2024, road fatalities in the European Union decreased by 3%, totaling around 19,800 deaths, but advancements differ across member states (Reuters, 2024).

Drowsy driver detection model significant to mitigate road accidents. Recently, the drowsy-driving-detection system has become a hot research topic in the field of computer science and engineering. The detection approaches are categorized as subjective and objective detection.

Deep learning is a sub-field of machine learning where the algorithm is inspired by the structure of the human brain. Examples of deep learning algorithms include but are not limited to, Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory networks (LSTMS), Deep Belief Networks (DBN), Generative Adversarial Networks (GAN).

Recently, many problems, such as image classification problems including drowsiness detection, have been tackled using deep learning-based approaches. In this study, ResNet 50 and DLIB shape predictor facial landmark model will be employed to develop a drowsy driver detection system.

Drowsy drivers are responsible for a large percentage of traffic fatalities and injuries worldwide, making road accidents a major public safety concern. A driver's attention, reaction time, and capacity to make decisions are all hampered by sleepy driving, which frequently has serious or deadly repercussions. Conventional techniques for identifying sleepiness, such self-reporting or physiological sensors (like EEG and heart rate monitors), are either invasive, unreliable in real-time, or too difficult to install in standard cars. A non-intrusive, precise, and real-time driver monitoring system that can successfully identify sleepiness indicators before they cause collisions is desperately needed. Although earlier systems have made an effort to use facial analysis for this goal, many of them are not robust because of poor generalization in lighting, occlusion, and position variation. Furthermore, depending just on one characteristic, such eye closure, can result in limited reliability and false positives.

In order to solve the issue, this study created a hybrid driver drowsiness detection system that combines two complementary methods: Dlib's facial landmark shape predictor with OpenCV-based head pose estimation to evaluate driver orientation, and ResNet-50 deep convolutional neural network for predicting drowsy or non-drowsy driver images. The method intends to improve overall robustness in a variety of real-world driving scenarios, decrease false positives, and increase drowsiness detection accuracy by combining these two models.

The drowsiness detection system is used in vehicle safety technology (VST). The findings of this study will assist researchers and engineers in the domain of VST in developing a drowsiness detection system that identifies early indicators of fatigue before the driver entirely loses focus and alerts them accordingly.

Related Works

Drowsiness is a major contributing factor to traffic accidents worldwide. Researchers are utilizing a variety of methods to develop driver drowsiness detection systems. Nevertheless, most research on sleepiness detection uses single-parameter approaches, which are insufficiently reliable and valid to be used in cars. When training a deep neural network, datasets are essential. When datasets are not represented, trained models are prone to bias because they are unable to generalize to real-world scenarios. It is particularly challenging to generalize models that were trained in certain cultural contexts since they might not adequately depict a wide variety of races. Due to the fact that many publicly available databases do not fairly represent all ethnic groups, this presents a special challenge for the detection of driver drowsiness.

Traditional augmentation strategies cannot improve a model's performance when tested on several groups with varying facial traits, and it is sometimes challenging to produce new, more representative datasets. A unique approach that improves driver sleepiness detection performance for various ethnic groups was presented by Ngxande et al. (2020). By employing Generative Adversarial networks (GAN) for targeted data augmentation based on a population bias visualization method that combines faces with similar facial traits and emphasizes areas where the model is failing, the framework enhances Convolutional Neural Networks (CNN) trained for prediction. After testing, the designed system's accuracy was found to be 96.91%.

Recently, deep convolutional neural networks have been used for a number of applications, including the identification of sleepy driving. Few studies have focused on the methodical adjustment of convolutional neural networks' (CNNs) hyperparameters, which could enhance the identification of driver fatigue. In order to close this gap, Faisal et al. (2022) introduced a deep learning technique to predict drivers' tiredness and offer an alerting system to stop accidents caused by drowsiness. 97.98% accuracy was recorded by the model.

Deng and Wu (2019) implemented Dri-Care, which uses video images to identify drivers' levels of weariness, including yawning, blinking, and eye closure time, without requiring them to wear any gadgets. A new face-tracking algorithm was developed to increase tracking accuracy because of the flaws in earlier algorithms. Additionally, a novel facial region identification technique based on 68 critical points was created. The drivers' condition was then assessed using these facial regions. Dri-Care can notify the driver with a fatigue warning by integrating the features of the mouth and eyes. According to the experimental results, Dri-Care's accuracy was about 92%.

While some are intrusive, may divert the driver, or need expensive sensors, most traditional methods for determining intoxication rely on behavioral characteristics. Consequently, a real-time, lightweight system for identifying driver fatigue was presented by Mehta et al. (2020) and implemented as an Android application. The system records the videos and locates the driver's face in each frame using image processing algorithms. Based on adaptive thresholding, the system identified facial landmarks and calculated the Eye Aspect Ratio (EAR) and Eye Closure Ratio (ECR) to determine the driver's level of fatigue. The model underwent testing and assessment. The accuracy was 84%, according to the results.

An embedded feature selection approach was proposed by Bencsik et al. (2023) and can be used as a building block in the system development of a neural network-based sleepiness detector. Because the Feature Prune Layer is positioned ahead of the first layer in the architecture, its weights vary according to the significance of the relevant input features and are removed repeatedly until the required number is attained. EEG data was used to test the algorithm since, according to the literature, it is one of the best markers of tiredness.

The algorithm was able to reduce the original feature set by 95% with just a 1% decrease in precision, according to the results. When the top 10% and top 20% of the initial features were chosen, the precision increased by 1.5% and 2.7%, respectively. Furthermore, when the original feature set is reduced by 95%, the suggested method performs better than the widely used Principal Component Analysis and the Chi-squared test, achieving 24.3% and 3.2% higher precision, respectively.

A CNN model for classifying ocular states is proposed by Jahan et al. (2023). Three CNN models were tested: 4D, VGG16, and VGG19. To identify tiredness based on eye state, a new CNN model called the 4D model was created. The model was trained using the MRL Eye dataset. The 4D model performed exceptionally well when trained using training samples from the same dataset (about 97.53% accuracy for predicting the eye condition in the test dataset). Two more pretrained models (VGG16, VGG19) performed worse than the 4D model.

This paper describes how to develop a comprehensive drowsiness detection system that anticipates a driver's eye condition in order to assess the driver's level of sleepiness and notify the driver prior to any serious risks to road safety.A hybrid CNN-LSTM model was presented by Chen et al. (2024) to capture temporal and spatial dependencies in eye and mouth activity. Eye aspect ratio (EAR) and mouth aspect ratio (MAR), which are essential for gauging tiredness, may be precisely extracted thanks to the Dlib68-point facial landmark detector. In a similar vein, Kumar and Adepoju (2025) used OpenCV for continuous face tracking in different lighting conditions and ResNet50 for eye-state categorization.

The use of EEG, ECG, EMG, and heart rate variability to identify mental weariness has been highlighted in recent research. It is now possible to obtain these signals in actual driving situations because to wearable technology and brain-computer interfaces (BCIs).In a noteworthy work, Zhang et al. (2025) used the Muse headband to create a lightweight EEG-based system that was processed using a 1D CNN classifier to detect early indicators of weariness with 93.4% accuracy.

The use of behavioral indicators, such as steering pattern, lane departure, and braking behavior, to identify drowsiness is a developing field. Usually, time-series analysis and rule-based decision logic are used in these models. Albadawi et al. (2023) created a hybrid framework for real-time fatigue classification using steering wheel angle and acceleration data processed by LSTM networks. Such versions can be deployed more easily because to integration with CAN-bus systems in contemporary cars. A new system that uses DLIB’s facial landmark predictor for head posture and a ResNet-50 classifier to identify drowsy drivers accurately by assessing facial features like head tilt and eye closure, proving effective in different lighting and angles was introduced in this paper.

Methodology

The methods used in this investigation is presented in this chapter. The suggested model, data collecting, resizing, data normalization, Resnet 50, technique of analysis, and performance evaluationare discussed under this heading.

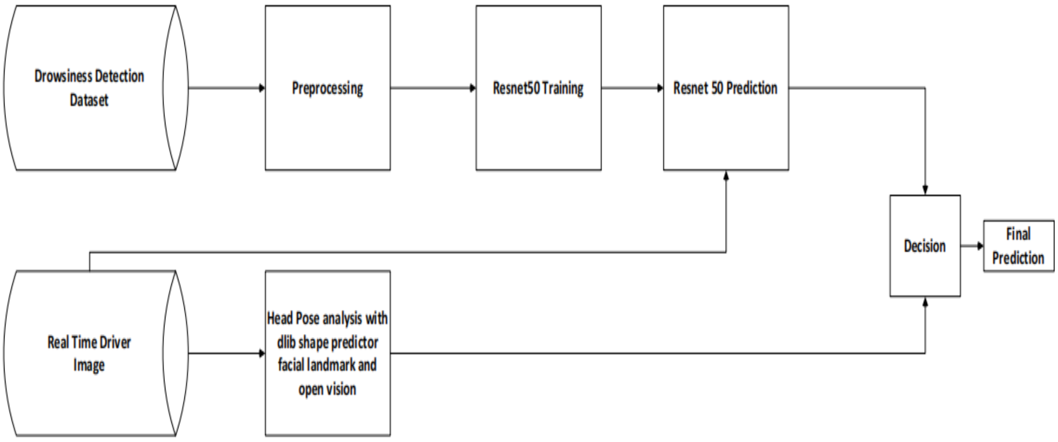

Fig 1: ***Structure of Drowsy Driver Detection***

The Proposed Model

The Figure 1 shows an organized process for a driver drowsiness detection system that combines two essential elements: head posture analysis using Dlib shape predictor and OpenCV and sleepy and non-drowsy prediction with ResNet-50. Two primary pipelines comprise the process: one for model training and another for real-time prediction. A Drowsiness Detection Dataset, which includes annotated photos of drivers classified as drowsy and non-drowsy, is the starting point of the first pipeline. In order to prepare the data for model training, this dataset goes through a preprocessing phase that usually entails actions like image scaling, normalization, and augmentation. A ResNet-50 model, a deep convolutional neural network ideal for image classification applications, is subsequently trained using the processed images. The model can predict eye states from fresh input photos once it has been trained.

Real-time driver monitoring is handled by the second pipeline. Two parallel analyses are carried out when a real-time driver image is taken. To determine if the driver is sleepy or not, the image is first run through the trained ResNet-50 model. Second, the same image is examined for head pose estimation using OpenCV to ascertain the driver's head position (e.g., looking straight, slanted, or nodding) and dlib shape predictor for face landmarks. A more accurate evaluation of drowsiness is made possible by this dual study.

A decision-making module receives the results of the head position analysis and the ResNet-50 prediction. This module uses rule-based or logical fusion. If head position analysis gives a value greater than 95 degrees for multiple frames and the output from ResNet50 is sleepy, drowsiness is expected. Overall, this hybrid technique creates a reliable and real-time drowsiness detection system by utilizing the advantages of deep learning for visual classification and classical computer vision for position estimation.

Data Collection

The sleepiness detection dataset from Kaggle (https://www.kaggle.com/datasets/starryfu/driver-fatigue-detection) was analyzed in this study. The faces of drivers from the Real-Life Drowsiness Dataset's videos were removed and cropped to create the Driver Drowsiness Dataset. VLC software was used to extract the frames as pictures from videos. There are two classes in the dataset: active subject and weary subject. The collection has 9120 photos in total, with 4560 images per class. Figure 2 displays a random sample of the dataset.

Figure 2: *Drowsiness Detection Dataset*

Resizing

The input images for the system were three colour channels (RGB) of a fixed size of 244 by 244 pixels in order to enable consistent processing. The resizing ensures that the spatial dimensions of the image match the model's input layer setup.

Data Normalization

The aim of normalization is to scale the input data to uniform range or distribution in order to generate numerical stability, accelerate model's convergence, and to enhance its performance during training and inference. The normalization that was based on the statistics of the dataset and mean subtraction was employed, and this was carried out automatically by the kera's ResNet50 model's preprocess-input () functionto understand input images.

ResNet 50

Pretrained Resnet 50 will be altered for artist classification in this investigation. The model will be modified by adding layers. Each of the five phases of Resnet 50 has an identification block and convolution. With one exception, each convolutional and identity block consists of three convolutional layers. This investigation will include an additional step. Figure 3 displays the suggested modification's architectural diagram.

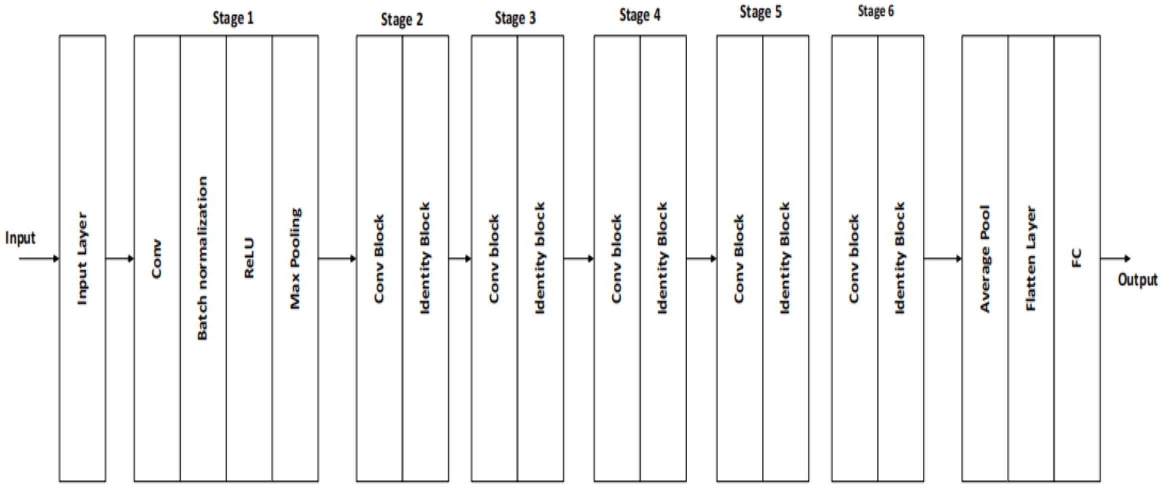

Figure 3*: Modified Resnet50 Architecture*

Residue Block

Exploding and disappearing gradients appear when CNN's layer count rises. To solve this issue, the residue block is used in the Resnet design. Given a neural network with input x, it is crucial to determine the true distribution H(x).

The residue can be represented as

R(x) = output – input = H(x) –x (3.1)

(Gao et al., 2023)

Which can be re-written as

H(x) = R(x) + x(3.2)

(Gao et al., 2023)

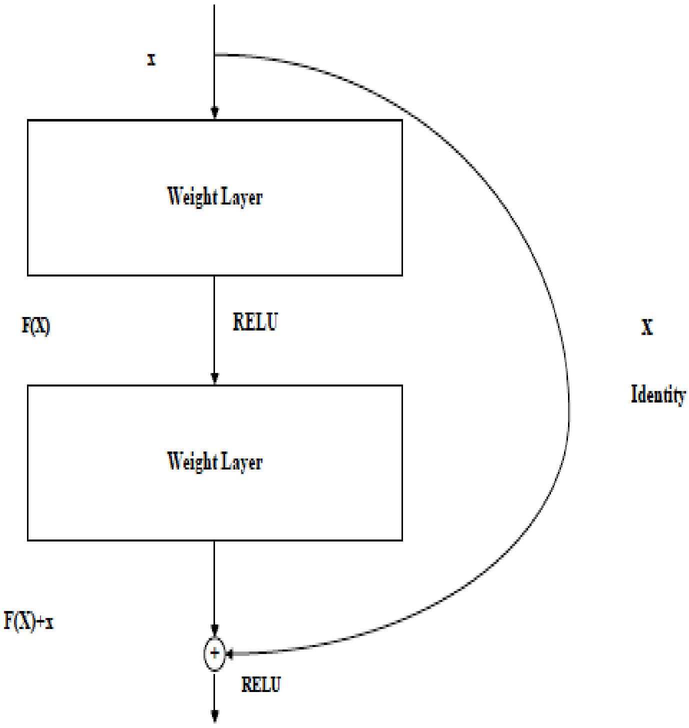

Figure 4: ***Residual Block***

The residue block is trying to figure out what the actual output is, H(x).Figure 4 shows that because x is causing an identity link, the layers are learning the residual, R(x). The layers in a standard network learn the true output (H(x)), but the layers in a residual network learn the residual (R(x)). Furthermore, it has been discovered that learning the residual of the input and output is easier than learning the input by itself. The identity residual model allows activation functions from previous levels to be reused in this fashion because they are omitted and do not complicate the design.

Head Pose Analysis with Dlib Shape Predictor Facial Landmark and Opencv

Head position analysis was determined using the Dlib shape predictor _facial _landmark model and open CV. This method uses 2D camera images to assess the head's orientation in 3D space, including whether the driver is tilting their head, gazing straight ahead, or nodding asleep. Face detection, which is often done with a pre-trained face detector, was the first step in the process. Dlib's 68-point facial landmark predictor was used after the face has been identified. Important face characteristics like the corners of the eyes, the tip of the nose, the chin, and the margins of the lips were located by this shape predictor. Because they serve as reliable anatomical reference points, these landmark sites are essential for evaluating head posture.

The Perspective-n-Point (PnP) algorithm was utilized in OpenCV to carry out the real head posture estimate following the acquisition of the face landmarks. In order to do this, a series of 2D facial points from the image was mapped to match 3D model points of a typical human face. Points like the corners of the eyes, the chin, and the tip of the nose, for instance, were mapped to fixed coordinates in three dimensions. The rotation and translation vectors that connected the 3D model to the camera's coordinate system were then estimated using OpenCV's solution PnP () function.

The pitch (up/down), yaw (left/right), and roll (tilt) of the head were calculated by interpreting these rotation and translation vectors. The driver's attention state can then be categorized using thresholds. For example, quick side-to-side movements may indicate distraction, whereas a head that is constantly inclined downward may indicate sleepiness or drifting off.

Method of Analysis

The analysis makes use of Open Vision Library, Dlib, Keras Framework, and Anaconda IDE programs. Stream-lit was also used in the design of the GUI.

Performance Evaluation

The developed model was evaluated in terms of accuracy, precision, recall, F1-score, and loss. The mathematical formular are as given below.

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \qquad (3.3)$$

$$Precision = \frac{TP}{TP+FP} \qquad (3.4)$$

$$Recall = \frac{TP}{TP+FN} \qquad (3.5)$$

$$F1\text{˜}score = \frac{2\text{â̂—}Precision\text{â̂—}Recall}{Precision+Recall} \qquad (3.6)$$

Where TP denote True Positive, TN represent True Negative, FN is the number of False Negative and FP represent False Positive.

Result and Discussion

The analysis conducted for this study is presented in this chapter. Data visualization, preprocessing, Resnet 50 implementation, head posture analysis using the Dlib shape predictor model, performance evaluation, and discussion are all included in the results.

Dataset Visualization

Figure 5 displays a random sampling of the dataset used in this investigation. Head posture, mouth position, and eye condition are among the details recorded in this dataset. These three details are sufficient to identify sleepy drivers.

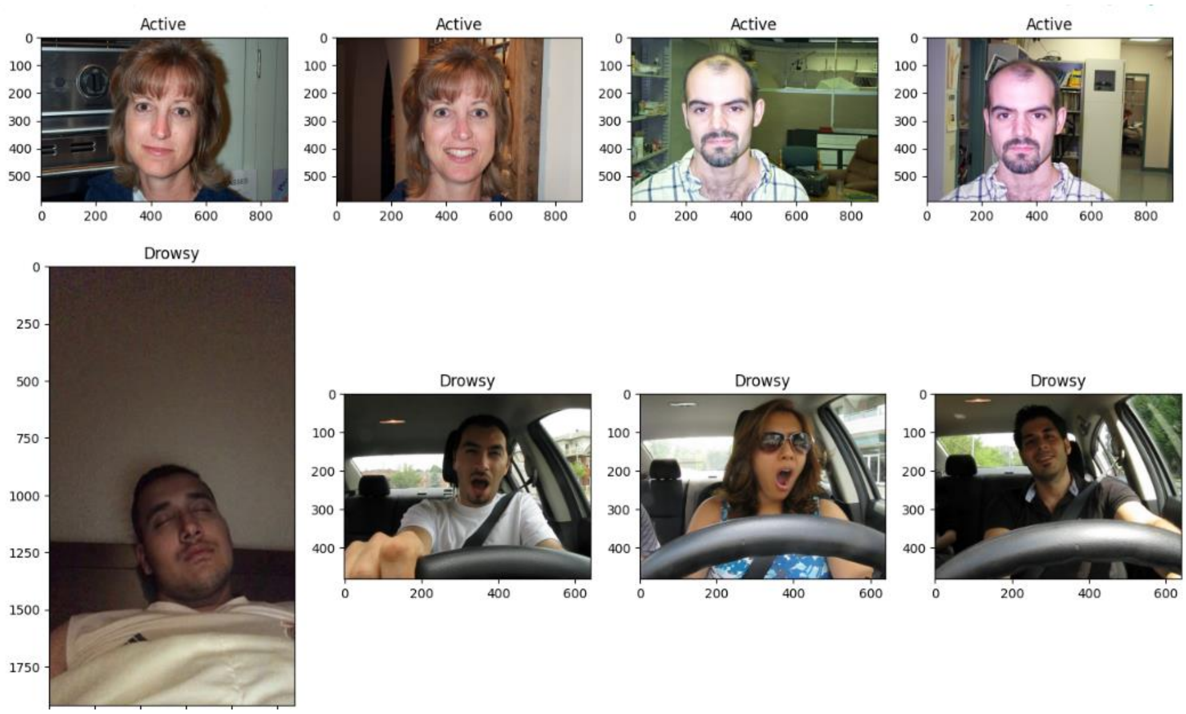

Figure 5: ***Random Sample of Dataset***

Prior to being fed into the ResNet50 model, the pictures underwent preprocessing. Resizing was the initial preprocessing step carried out. ResNet50 was used to take input images with three color channels (RGB) and a fixed size of 224 by 224 pixels. Since the dataset's photos vary in size, they were all downsized to 224 by 224 so that the model could process them effectively. Normalization, which converts pixel values to a range and distribution that the model anticipates based on its initial training data, is another crucial phase carried out. When ResNet50 is pretrained on ImageNet, it is assumed that input pixel values are first normalized channel-wise using the following mean and standard deviation values after being scaled to a [0, 1] range by dividing by 255. This suggests that by deducting the mean and dividing by the standard deviation, each color channel (Red, Green, and Blue) was separately normalized. This stage stabilizes predictions during inference and aids in accelerating convergence during training.

Implementation of ResNet 50

Tensor flow and the Keras deep learning Python package were used to construct ResNet 50. The hyperparameters listed in Table 1 were used for the implementation. The dataset was divided into training, testing, and validation groups in a 70:20:10 ratio respectively. Hyperparameters from tuning 1 and tuning 2 were used to train the model. The final choice includes values that produce the best results.

Table 1: ***Hyperparameter Settings***

	Hyperparameters	Tuning 1	Tuning 2	Final Selection
1	Activation function (input layer)	Rectified Linear Unit	Rectified Linear Unit	Rectified Linear Unit
2	Activation function output layer	Softmax	Softmax	Softmax
3	Loss function	Sparse categorical cross entropy	Mean square error	Categorical cross entropy
4	Batch size	32	16	64
5	Number of Epoch	5	50	100

The model's performance across 50 epochs is shown in the training and validation accuracy plot. The model quickly picks up significant patterns from the data, as evidenced by the initial rapid increases in both training and validation accuracy. The curves stabilize after the initial few epochs, with both metrics getting closer to values near 0.9. This implies good learning and significant convergence. The training and validation accuracy lines stay closely aligned throughout the training phase, which is a sign of good generalization. The small difference between the two curves suggests that the model performs consistently on both seen and unseen data and is not severely affected by overfitting. Although they don't show any instability, little changes in validation accuracy are expected because of differences in the data batches. All things considered, the plot shows a strong training procedure with outstanding performance results.

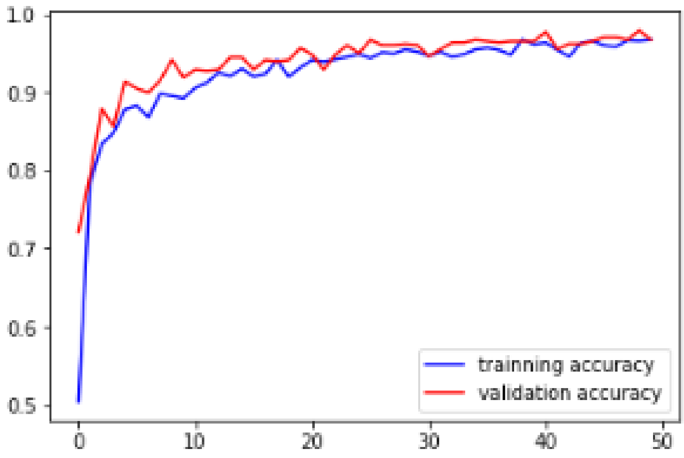

Figure 6: ***Training History for Accuracy***

The model's learning behavior across 50 epochs is depicted in the training and validation loss plot. Both the training and validation losses are comparatively high at the beginning of training, with the training loss starting at more than 1.0. But during the first few epochs, they drastically drop, suggesting that the model is rapidly reducing error and picking up useful representations. Both charts show high performance and low prediction error as training goes on, declining steadily until they converge to very low levels close to 0.05 by the end of the epochs. The training and validation loss curves closely overlap, indicating that the model is not overfitting and has good generalization.

In order to get strong forecast accuracy on unknown data, this balance is essential. Both curves exhibit slight variations, which are normal and do not suggest any instability in the training procedure. With steadily declining loss and no indications of divergence between training and validation performance, the plot generally shows an efficient training procedure.

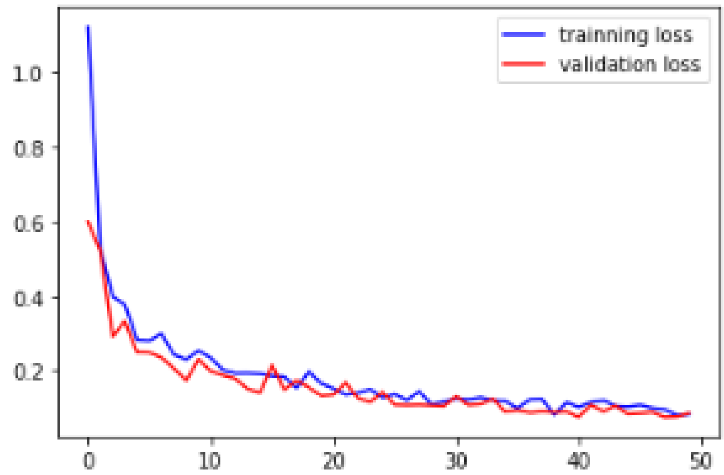

Figure 7: ***Training History for Loss***

Table 2: ***Performance Evaluation***

Accuracy	Precision	Recall	F1-score
90.0%	0.88	0.88	0.88

Implementation of Shape Predictor Face Landmark Model from Dlib

A drowsy driver detection system, especially one based on facial indicators like eye closure or head position, depends heavily on the use of the shape predictor face landmark model from DLIB. The first step in the procedure is to use OpenCV to record live video frames. The frontal face detector in DLIB analyzes every frame and determines the coordinates of any faces it detects. The shape_predictor_68_face_landmarks were used once a face is identified. Sixty eight (68) distinct facial landmarks on the identified face are mapped using the data model. Important face features like the eyes, nose, mouth, jawline, and eyebrows were represented by these landmarks.

The pitch, roll, and yaw of the head were estimated using landmarks on the nose tip, chin, eyes, and corners of the mouth. Drowsiness is indicated by a prolonged forward or side tilt of the head (high pitch angle).

Deployment into GUI Application

The head posture analysis-based drowsy driver detection system uses a webcam and real-time facial landmark monitoring to determine a driver's head orientation. To provide an interactive web-based interface, the system was constructed using Python, Keras, OpenCV, Dlib, Resnet50, and Stream-lit. The computer determines the pitch angle of the head by recognizing particular facial characteristics, such as the nose tip, chin, eyes, and mouth corners. The technology recognizes this as an indication of tiredness and shows a warning message on the screen if the driver's head tilts sideways more than a certain threshold for several consecutive frames. Prediction for ResNet50 and shape_predictor_68_face_landmarks predictor models were merged to provide the output for this application's successful execution. This technique offers a camera-based, non-intrusive way to track fatigue, which is particularly helpful for vehicle safety applications. It is accessible and easy to use for testing or demonstration because it operates in real time and can be deployed via a straightforward web interface utilizing Stream-lit. An example of detection carried out by the GUI application is shown in Figure 8.

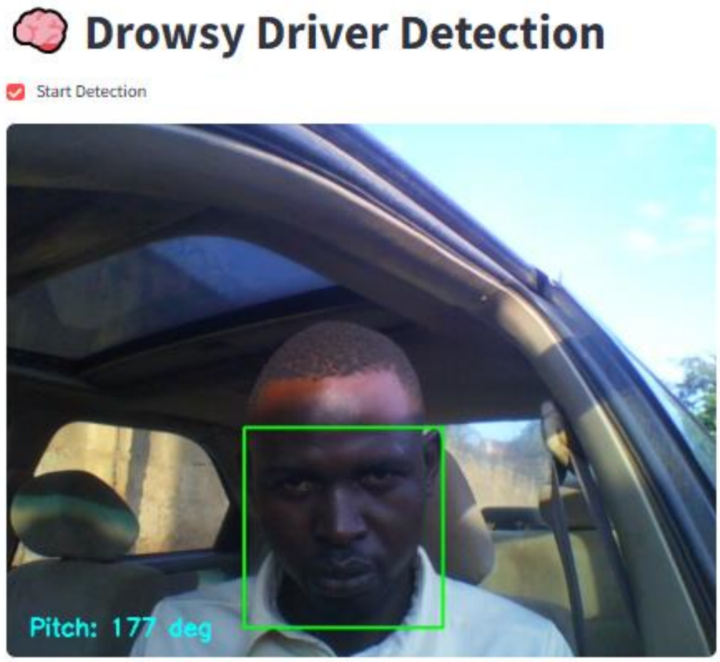

Figure 8: ***Drowsy Driver Detection Showing Non-Drowsy Driver***

As shown in Figure 8, the Dlib shape predictor analyzes if a head position is tilted over the predetermined threshold, whereas the Resnet 50 model predicts whether the driver is sleepy or not. Figure 9 makes it clear that the driver is sleepy, and the model was able to identify this.

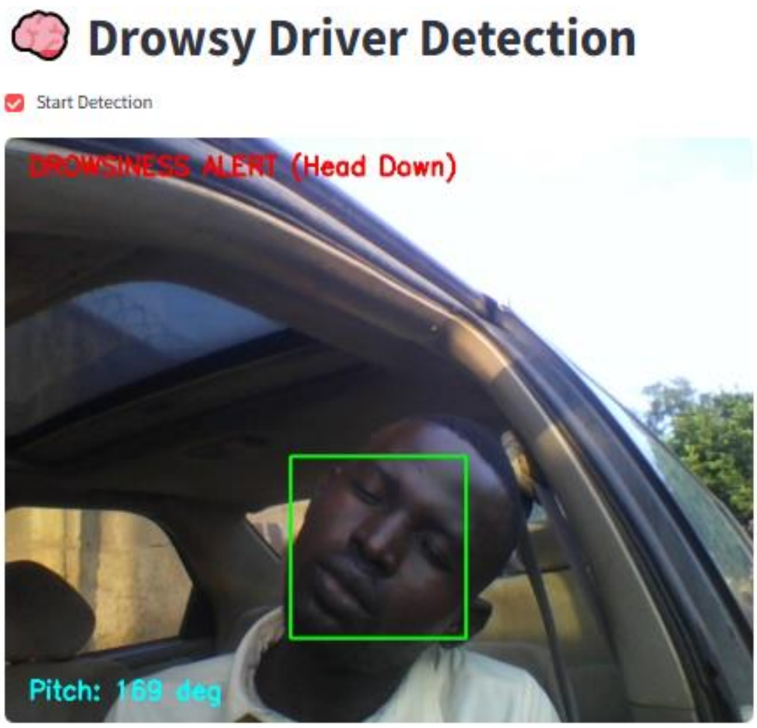

Figure 9: ***Drowsy Driver Detection GUI Showing Drowsiness***

Discussion

A significant proportion of traffic accidents and fatalities are caused by drowsy driving, which is still a major concern for road safety worldwide. In order to solve this issue, the suggested system combines the Dlib facial landmark shape predictor for head posture estimation with deep learning, notably ResNet-50 for drowsy and non-drowsy prediction. To determine if a facial image is sleepy or not, the ResNet-50 Convolutional Neural Network is utilized. By using deep residual learning to counteract vanishing gradients, this methodology enables the formation of extremely deep networks that perform well when applied to real-world data. Because of its excellent performance in picture classification tasks, its use in sleepiness detection is very advantageous. Even in difficult lighting situations or with occlusions like spectacles, ResNet-50 can consistently discern minute variations between open and closed eyes. The model's high accuracy makes it appropriate for real-time use. It was trained on a carefully selected dataset of labeled eye pictures.

The 68-point shape predictor from the Dlib library was used to extract facial landmarks in addition to ResNet50 prediction. Pitch, yaw, and roll—the three angles that indicate the driver's head orientation can be estimated thanks to these landmarks. Head motions like tilting sideways or downward over prolonged periods of time are thought to be signs of tiredness or exhaustion. This non-invasive method of estimating head posture is computationally effective and doesn't need any extra gear, such depth cameras or infrared sensors. The technique gets over the drawbacks of single-modality methods by combining head posture analysis with ResNet50 classification.

The technology can be integrated into commercial vehicle driver assistance systems because it is made for real-time deployment using a webcam and moderate processing power. Despite being deep, the ResNet-50 model can be improved with Tensor RT or transformed into a Tensor Flow Lite model for edge devices. The Dlib library maintains minimal latency for facial landmark detection since it is lightweight and developed in C++. Nevertheless, head movement outside the camera's field of view, occlusion from sunglasses, and dim lighting continue to pose problems for real-time deployment. Future research may use sensor fusion or infrared imaging to overcome these constraints.

Conclusion and Recommendation

Conclusion

In this work, a hybrid drowsiness detection system was created by combining a ResNet-50-based classifier for identifying driver photos as drowsy or non-drowsy with DLIB's 68-point facial landmark predictor for head posture analysis. A reliable assessment of driver alertness is made possible by the model's accurate capture of crucial facial characteristics that indicate weariness, such as head tilt and prolonged eye closure. The system demonstrated its potential application in real-time vehicular safety systems by successfully detecting drowsiness under a variety of lighting and orientation conditions by combining geometric-based and deep learning-based techniques.

Recommendations

(i) Real-World Testing: Extensive field testing in real driving environments is necessary to validate and fine-tune the system for deployment.

(ii) Driver Feedback Integration: Implement alert mechanisms (audio, haptic, or visual) and assess driver responsiveness to warnings.

(iii) Collaboration with Automotive Industry: Collaborate with automotive manufacturers for integration and testing in advanced driver assistance systems (ADAS).

(iv) Ethical and Privacy Considerations: Ensure compliance with data privacy standards, particularly when handling facial imagery in public or shared vehicles.

References

Albadawi, Y., Alredhael, A. & Takuri, M. (2023). Real-Time machine learning based driver drowsiness detection using visual features. J IMAGING, 9(5), 91. https://doi.org/10.3390/jimaging9050091

Bencsik, B., Remenyi, I., Szemenyel, M., & Botzheim, J. (2023). Designing an embedded feature selection algorithm for a drowsiness detector model based on electroencephalogram data. *Sensors*, *23*(4). https: lldoi.org\10.3390\s2304-1874.

Chen, X., Qu, S., Gao, Z., Li, N., Wang, Y., & Wu, X. (2024). Multi-task learning for fatigue detection and face recognition of drivers via tree-style space-channel attention fusion network. *arXiv preprint arXiv:2405.07845*.

Deng & wu, (2019). Deep-learning based head pose estimation for driver monitoring systems. *International journal of computer vision*, *127*(8), 1234-1248.

European Commission (2025). EU Road Fatalities Drop by 3% in 2024, But Progress Remains Slow. https://transport.ec.europa.eu/news-events/news/eu-road-fatalities-drop-3-2024-progress-remains-slow-2025-03-18_en

Faisal, T., Negassi, I., Goitom, G., Yassin, M., Bashir, A., & Awawdeh, M. (2022). Systematic development of real-time driver drowsiness detection system using deep learning.

Ghimire, D., Jeong, S., Yoon, S., Park, S., & Choi, J. (2025). Real-time sleepiness detection for driver state monitoring system.

Jahan, I., Uddin, K. M., Murad, S. A., Miah, M., Khan, T. Z., Masud, M., & Bairagi, A. K. (2023). 4D: A real-time driver drowsiness detector using deep learning. *Electronics*, *12*(1), 235.

Mehta, S., Dadhich, S., Gumber, S., & Bhatt, A. J. (n.d.). *Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio Fatigue Detection Non-Intrusive Methods Driver monitoring system*.

Ngxande, M., Tapamo, J. R., & Burke, M. (2020). Bias remediation in driver drowsiness detection systems using generative adversarial networks. *IEEE Access*, *8*, 55592–55601.

World Health Organization (2024), Ro*ad Safety*. https://www.who.int/health-topics/road-safety.

World Health Organization (2024).*Road Traffic Injuries*. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Zhou, C., Zhao, Y., Liu, S., Zhao, Y., Li, X., & Cheng, C. (2024, May). Research on driver facial fatigue detection based on Yolov8 model. In 2024 5th International Conference on Information Science, Parallel and Distributed Systems (ISPDS) (pp. 282-285). IEEE.