ML-based Eye-Blinking Tracking for Fatigue Detection using Visual Data

Arpita Singh[1] · **Shruti Bhargava**[1] · **Snehil Sharma**[1]

Abstract

Drowsiness-related fatigue is a major cause of road accidents that demand constant focus. Driver fatigue accounts for around 1 in 5 traffic accidents. Thus, fatigue is an indicator of baseline risk for the occurrence of errors and accidents. The eye-blinking-based method is a promising approach to detect driver drowsiness. This process involves monitoring the pattern and frequency of eye-blinks while driving. Our system utilizes a CNN model, an efficient deep learning architecture, to accurately classify driver drowsiness from real-time video feeds. We enhance the system's accuracy with the Haar Cascade algorithm, which efficiently detects facial features. Furthermore, the proposed method allowed a real-time eye rate analysis, where the threshold served as a separator of the eye into two classes, the "Active" and "Sleepy" states. By predicting the driver's eye state based on these categories, the system determines the driver's drowsy state and alerts the driver before any severe threats to road safety.

Keywords Fatigue · Driver · Eye · Tracking · Alert · CNN · OpenCV

✉ Arpita Singh
arpitasingh.rajput16@gmail.com

Shruti Bhargava
shrutibhargava2004@gmail.com

Snehil Sharma
ssnehil06@gmail.com

Supervisor(s)

Prof. Sanjay Garg[1]
sanjay.garg@juetguna.in

Dr. Amit Kumar[1]
amit.kumar@juetguna.in

[1] Department of CSE, Jaypee University of Engineering and Technology, Guna, MP, India

1 Introduction

Machine learning help systems learn and adapt like humans. Instead of relying on fixed rules, it looks for patterns to figure out if driver is getting tired.

By Combining advanced image processing techniques with machine learning algorithms, our system can accurately detect and analyse subtle eye closure patterns, which are crucial indicators of fatigue. This innovative approach allows the system to adapt to individual differences in eye movements, ensuring both precision and reliability. Using visual data from cameras or video streams, the system allows for real-time monitoring of eye activity.

In an operational environment scenario where, operating systems are dependent on human performance, fatigue can be defined as an inclination to degrade performance. To address this problem, we propose a real-time drowsiness detection system utilizing deep learning algorithms, specifically Convolutional Neural Networks (CNNs), that emerged as a powerful tool for drowsiness detection. By automatically learning from input data, CNNs can effectively extract key features directly from raw visual information, such as eye-blink patterns.

The proposed system operates by capturing video feeds of the driver's face and analysing the frames for signs of drowsiness. The Haar Cascade classifier first detects the face and eyes, ensuring that the CNN processes only relevant regions of interest. If drowsiness is detected, the system can trigger alerts, such as auditory warnings to prompt the driver to take necessary actions.

The Social Impact By using cutting-edge machine learning technology, we can detect the early signs of driver fatigue, providing real-time alerts to prevent accidents before they happen. This technology not only helps save lives but also contributes to making our roads safer and less stressful for everyone, empowering drivers to stay alert and focused on the road ahead.

Role of Technology in Combating Fatigue Detection Driver fatigue is one of the leading causes of road accidents, especially on long trips or during late hours. Thanks to advancements in technology, we now have tools that can help keep drivers safe by detecting drowsiness before it becomes dangerous. Modern systems use smart technologies like AI to watch for signs of fatigue. For example, cameras can track how often a driver blinks or if their eyes are closing too often. These tools are making a real difference by helping drivers stay alert and preventing accidents caused by drowsiness.

The Purpose of the Research The primary purpose of the driver drowsiness detection system is to continuously monitor the driver's alertness by tracking both active and sleepy states in real-time and triggering an alarm to alert the driver when drowsiness is detected. Ultimately, this system aims to utilize advanced technology, promote awareness and enhance road safety.

The subsequent sections of this paper are structured as follows: In Sect. 2, we provide a concise overview of related works. Section 3 outlines the methodology employed for drowsiness detection. In Sect. 4, you'll find details about the dataset and the results. Lastly, in Sect. 5, we offer conclusions and identify potential directions for future research in this domain.

2 Literature Survey

The analysis of driver drowsiness detection is an emerging field that is gaining prominence due to its critical implications for road safety and accident prevention. In this section, we review several works in this domain.

Safarov et al. [1] introduced a comprehensive approach to driver drowsiness detection by integrating deep learning techniques with computer vision. Their method employed a custom dataset that captured a range of driver states, including awake-open eyes, drowsy-closed eyes, and yawning, achieving a remarkable accuracy of 95.8% for detecting drowsy eyes and 97% for open eyes. The researchers utilized facial landmark detection to analyze eye-blink patterns and mouth movements, establishing a correlation between yawning and drowsiness. Despite these advancements, the study noted significant challenges, particularly in scenarios where drivers wore sunglasses, which obstructed the detection of eye landmarks and adversely affected the system's performance. This limitation underscores the need for further refinement in the model to enhance its robustness against real-world variations and improve its applicability in diverse driving conditions.

Zhang et al. [4] developed an advanced driver drowsiness detection system utilizing a multi-feature fusion approach that integrates facial landmarks with deep learning techniques. Their research involved analyzing various datasets to assess driver states under different conditions, achieving an impressive accuracy rate of 95.1% in detecting fatigue levels. The study emphasized the importance of combining multiple facial features, such as eye and mouth states, to enhance the reliability of drowsiness detection. However, despite these advancements, their model encountered limitations in real-time processing speeds, particularly in dynamic driving environments. The authors noted challenges related to the temporal assembly process during data integration, which resulted in increased memory usage and processing overhead. This aspect highlighted the necessity for optimizing the model to balance accuracy with computational efficiency, particularly for deployment in real-time systems where rapid decision-making is crucial for driver safety.

Smith et al. [5] proposed an innovative approach to emotion recognition in videos by leveraging deep learning techniques and temporal convolutional networks. Their research focused on analyzing emotional expressions across diverse video datasets, which included both scripted and spontaneous interactions. By employing a multi-stream architecture, they successfully captured the nuances of emotional transitions over time, resulting in a classification accuracy of 92% for detecting six basic emotions. Despite these promising results, their model encountered significant challenges related to computational efficiency, particularly when processing high-resolution video streams in real-time. The need for extensive computational resources limited the model's applicability in resource-constrained environments, such as mobile devices. Additionally, while the model demonstrated robustness in controlled settings, it struggled to generalize effectively to unstructured, real-world scenarios where emotional expressions may be subtle or ambiguous. These limitations underscore the necessity for further research to enhance the model's adaptability and efficiency in practical applications.

Ambak et al. [9] conducted a comprehensive review of Intelligent Transport Systems (ITS) aimed at enhancing motorcycle safety, highlighting their significant potential in improving traffic safety for this vulnerable group of road users. The authors explored various existing and emerging ITS technologies, including advanced driver assistance systems (ADAS), collision warning and avoidance systems, lane-keeping and lane-change warning systems, intelligent speed adaptation, and visibility-enhancing systems. They noted that while many ITS applications have been developed for cars and commercial vehicles, there is a notable lack of systems specifically designed for motorcycles, which poses unique challenges due to their distinct characteristics and safety needs.

3 Methodology

We have presented a method to enhance the accuracy of detecting driver drowsiness by analysing eye blinking patterns, providing real-time alerts to prevent road accidents. The implementation was completed through a five-phase process called dataset, dataset pre-processing, model training, model testing and evaluation, the detailed structure of which is illustrated in Fig.1.

Real-Time Detection Using Haar Cascade: For real-time implementation, the Haar Cascade algorithm was used to detect and localize the driver's face and eyes within video frames. This method provides a fast and efficient way to isolate facial regions, allowing the CNN model to focus on the eye area for drowsiness classification. Once the eyes are detected, each frame is processed to identify whether the eyes are open or closed, determining the driver's active or sleepy state.

3.1 Data Set Description

This paper makes use of a private dataset known as the drowsiness dataset. The dataset contains images of open and closed eyes (726 images in each) essential for training a model to recognize drowsiness.

More information can be found in the Results analysis section.

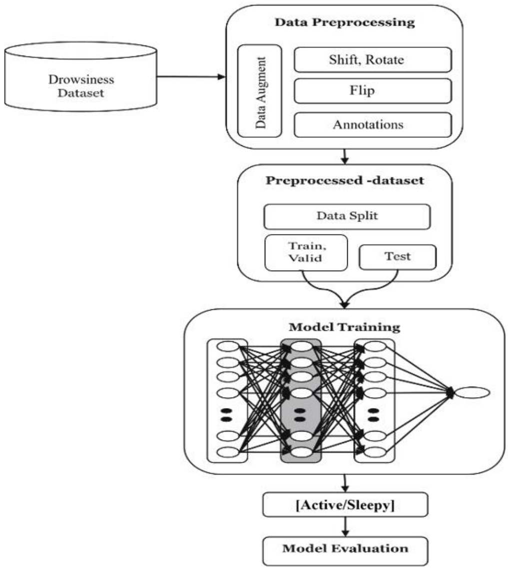

Fig. 1 Eye-blinking tracking for fatigue detection

3.2 Data Set Pre-processing

During the data pre-processing phase, we applied data augmentation such as shifting, rotating and flipping (Generated 5 augmented image per original image). We also resized the images and created an annotation file for labelling. Once the data was appropriately pre-processed, we proceeded to split it into 80% training, 10% validating and 10% testing subsets. This partitioning is crucial for assessing the model's performance and generalization on unseen data. The training data was used to train our model, while the test data was reserved to evaluate the model's predictive capabilities.

3.3 Model Training

During this phase of the project, we focused on training our model using the carefully prepared data. We chose to implement a Convolutional Neural Network (CNN) for classification tasks. CNNs are powerful deep learning models specifically designed to work with grid-like data, such as images. Their unique ability to learn and identify patterns and features in spatial hierarchies makes them particularly effective for tasks in computer vision, including image classification, object detection, and segmentation.

To optimize our model, we used the Adam optimizer with a learning rate of 0.0001. Since we were dealing with a binary classification problem, we selected 'binary_crossentropy' as our loss function, and we monitored the model's performance using accuracy as a key metric.

To enhance the model's performance and prevent overfitting, we implemented L2 regularization. This technique introduces a penalty to the loss function, discouraging the model from assigning excessively large weights to its parameters. By applying L2 regularization to both the convolutional and dense layers, we aimed to create a more robust model that could handle noise in the training data.

We also incorporated two important strategies during the training process: Early Stopping and Learning Rate Scheduling. Early Stopping keeps an eye on the validation loss, halting the training process if it doesn't show improvement for five consecutive epochs. This approach ensures that we retain the best weights achieved during training. Meanwhile, Learning Rate Scheduling helps us fine-tune the model by reducing the learning rate by 10% every 10 epochs, which aids in achieving effective convergence.

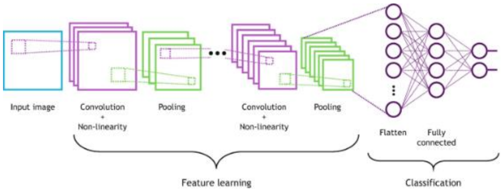

Fig. 2 CNN Architecture

3.4 Model Testing and Evaluation

In the realm of Fatigue Detection, model testing and evaluation is a critical phase that determines the effectiveness and reliability of the developed algorithms. The assessment process involves a comprehensive analysis of the model's performance using our dataset.

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

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

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

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

```
F1 Score :  94.47708578143362
Precision Score :  91.9908466819222
Recall Score :  97.10144927536231
```

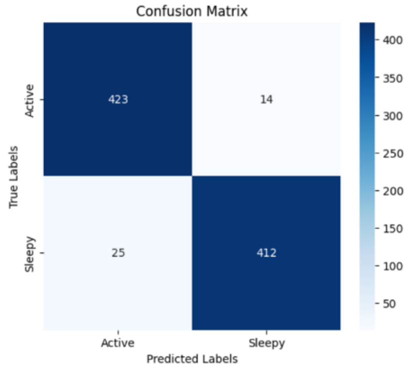

Fig. 3 Confusion matrix

4 Result Analysis

In this section, we took a closer look at some of the most important findings that may be used to gauge how well the model is working. These metrics assess how well our model can categorize or evaluate predictions. Sample instances of dataset are presented in Fig. 4.

Fig. 4 A snapshot of the data set

The training ROC curve displays the model's ability to distinguish between the two classes based on its predictions on the test data. From the Figure it is observed that, an AUC of 0.99 indicates that the model performs exceptionally well on the test set, as it achieves a high true positive rate while maintaining a low false positive rate across various classification thresholds. It is also observed that, a high AUC on the test data suggests that the model has effectively generalized and can reliably separate positive and negative examples in unseen data.

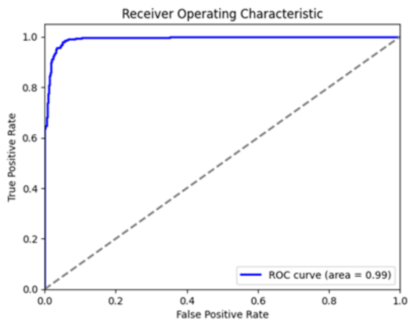

Fig. 5 ROC Curve

In summary, the model demonstrates a strong performance in distinguishing between the two classes, making accurate predictions.

5 Conclusions

The experimental results demonstrated the accuracy of proposed model to be 95.54%. As we strive for safer roads, the development and implementation of effective detection systems are essential. By leveraging technology to monitor driver alertness, we can significantly reduce the risks associated with drowsy driving, enhance road safety, and ultimately save lives. The ongoing research and innovation in this field will play a crucial role in shaping the future of transportation safety. We conclude that our study may be used as a key building block for assessing the credibility of fatigue detection in light of the acquired results.

Author Contributions Shruti Bhargava developed CNN model using python, methodology was written, and lead the team. Arpita Singh collected the data from Kaggle and preprocessing it. Snehil Sharma wrote the literature section and the Results.

Funding There was no funding for this research.

Availability of Data and Materials Data is available within this link https://www.kaggle.com/datasets/dheerajperumandla/drowsiness-dataset.

Code Availability The code was available with the authors. It will be provided on personal request.

Declarations

Conflict of interest The authors affirm that they do not have any competing interests.

Ethics Approval This material is the authors' own original work, which has not been previously published elsewhere.

Consent to Participate We all voluntarily agreed to take part in this study.

Consent for Publication We are agree with all the publication rules.

References

1. Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety by Furkat Safarov, Farkhod Akhmedov, Akmalbek Bobomirzaevich Abdusalomov, Rashid Nasimov, Young Im Cho. [MDPI]
2. Driver's Drowsiness Detection by A.Milan Hussain, Bhanu Sankar Ravi, Josaiah W. Lyngdoh N. [pdf]
3. Chand, H.V.; Karthikeyan, J. Cnn based driver drowsiness detection system using emotion analysis.(2022) *Intell. Autom. Soft Comput*. [Google Scholar]
4. Research on a Real-Time Driver Fatigue Detection Algorithm Based on Facial Video Sequences by Tianjun Zhu, Chuang Zhang, Tunglung Wu, Zhuang Ouyang, Houzhi Li, Xiaoxiang Na, Jianguo Liang and Weihao Li. [MDPI]
5. Eye Finding via Face Detection for a Foveated, Active Vision System by Brian Scassellati. [AAAI]
6. Eye Tracking based Driver Fatigue Monitoring and Warning System by Hardeep Singh. [Academia]

7. R. Grace, V.E Byrne, D.M Bierman, J.M Legrand, D Gricourt, R.K Davis, et al.(1998) A Drowsy Driver Detection System for Heavy Vehicles. [IEEE]
8. R. Dawn, "National Survey of Distracted and Drowsy Driving", 2003. [pdf]
9. K. Ambak, R Atiq and R Ismail, "Intelligent Transport System for Motorcycle Safety and Issues", Proceedings of the European Journal of Scientific Research, 2009. [Academia]