INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 09 ISSUE: 03 | MARCH - 2025 SJIF RATING: 8.586 ISSN: 2582-3930

Driver Drowsiness detection using Convolutional Neural Network

Satyam Kush[1]
Department of Computer Science and Engineering Chandigarh University,
Mohali, Punjab
strangekaisoul@gmail.com

***Abstract*— Driver drowsiness and tiredness are critical factors contributing to road accidents worldwide. Several accidents happen every year because of drowsy driving. This research paper introduces a pioneering approach to tackle driver drowsiness using convolutional neural networks (CNNs) based on the inception V3 architecture in automotive safety systems. The system focuses on monitoring the driver's facial cues, particularly eye movements, to identify instances of prolonged eye closure as indicators of potential drowsiness. The model is trained on a dataset comprising 84,898 images of open and closed eyes from 37 individuals. Training, validation, and test accuracies and losses, along with model architecture details, parameters, and performance metrics, is extensively analyzed and reported. The proposed system aims to enhance driver safety by issuing alerts when the driver exhibits prolonged eye closure, surpassing a predefined threshold. This study outlines the challenges, details the CNN-based model's design, dataset considerations, and showcases its robustness through real-world driving scenario experiments. Emphasizing the efficiency of real-time deployment, this study discusses optimizations and suggests future enhancements.**

Keywords—convolutional neural network (CNN), transfer learning, deep learning, driver state monitoring(DSM), InceptionV3

I. INTRODUCTION

Sleep is a dynamic process, and the inability to stay awake throughout activities is referred to as sleepiness. Sleep deprivation is characterized by insufficient sleep, either in quantity or quality. Reduced response time and poor decision-making skills are two examples of how sleepiness affects brain function. It is a major contributing factor to traffic accidents, which usually happen when a motorist is drowsy while driving a vehicle or because to abnormal sleep patterns, lack of sleep, drinking, or using prescription drugs. According to the National Safety Council (NSC), sleepy driving causes around 100,000 collisions, 71,000 injuries, and 1,550 fatalities annually [1].

In order to reduce the accidents due to driver drowsiness, a monitoring system is used to monitor the driver's facial expression and predict the driver's drowsiness level. The driver state monitoring (DSM) system alerts drivers when they become fatigued or drowsy. The primary camera sensor used by the DSM systems is mounted close to the dashboard and measures the pupil states, head positions, and eye blinks of the image. Eyelid closure metrics can be used to track and forecast driving impairment related to sleepiness (out-of-lane incidents)[2][3]. An illustration of a driver status monitoring system is shown in Figure 1, which alerts the driver by turning on the alarm, forcing the driver to pay attention to driving when it senses tiredness or inattention. [4].

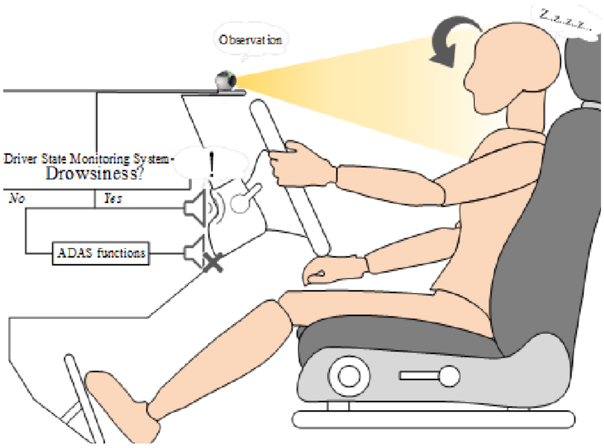

Fig. 1. Driver state monitoring system

The DSM system is based on a convolutional neural network. A specific kind of deep learning algorithm called a convolutional neural network (CNN) is used to process and analyse visual input, particularly photographs. It is composed of several layers that work together to automatically and adaptively learn hierarchical patterns and features from input images. CNN InceptionV3 architecture is used in the proposed model. InceptionV3 is a powerful and widely used CNN architecture known for its ability to capture multi-scale features efficiently, making it valuable in tasks requiring nuanced understanding of visual information [5].

II. METHODOLOGY

In this section , we discuss the benefits of using the Convolutional neural network(CNN) for the proposed model. A description of the dataset collection, preprocessing, and class distribution is presented. How the model is trained , what parameters and optimization techniques are employed, and how the scoring mechanism triggers an alert are explained here.

A. Convolutional Neural Network

One kind of deep learning neural network design that is frequently utilized in computer vision is the convolutional neural network (CNN). A computer can comprehend and analyze an image or other visual data thanks to the artificial intelligence discipline of computer vision. The input layer, convolutional layer, pooling layer, and fully connected layers are some of the layers that make up a convolutional neural network. [6].A simple CNN architecture is shown in figure 2.

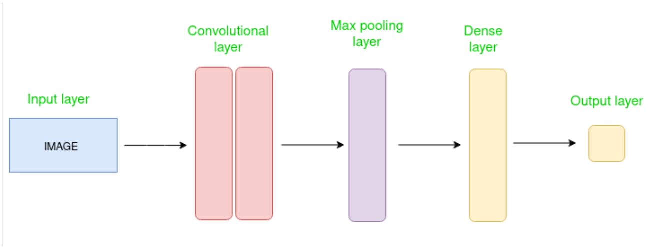

Fig. 2. Simple CNN architecture

We use transfer learning in the suggested model in place of training a CNN from start on a fresh dataset. The process of fine-tuning a pre-trained CNN on a fresh dataset is known as transfer learning. As a feature extractor, the trained CNN captures high-level visual representations. [7]. Through transfer learning, information from a machine learning model that has already been trained is applied to a separate but related problem[8]. Transfer learning architecture is shown in figure 3.

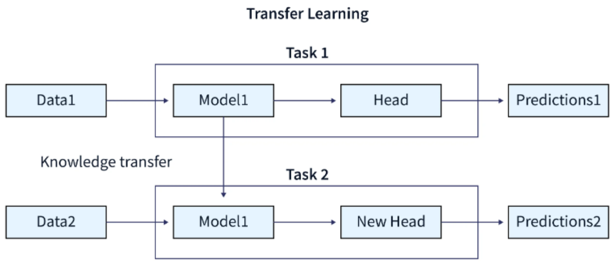

Fig. 3. Transfer learning architecture

In transfer learning, InceptionV3 is used in our model to save time and obtain accurate results . Over 78.1% accuracy in picture identification is achieved by the InceptionV3 model on the ImageNet dataset[9]. To extract features from photos, InceptionV3's pre-trained layers are employed.. Then, these features are fed into additional layers that are specific to the new classification task.

B. *Dataset Preparation*

Model trained on the dataset taken from the MRL Eye Dataset, which comprises 84,898 images of open and closed eyes from 37 different individuals. Images in the dataset are of different states including open eyes, closed eyes, good lightening, bad lightening, glasses over eyes, naked eyes, and reflection of flash over retina, thus making it a perfect dataset for model training [10]. Figure 4 contains the images from MRL eye dataset.

Fig. 4. Images from MRL eye dataset

The dataset is divided into two folders: the train and test folders. Each folder contains two classes: open and closed eyes. Train folder contains approximately 80% of the images of the dataset (i.e., 64,644 images), and the test folder consists of approximately 20% of the images of the dataset (16,160 images). Again Train folder is divided into two folders: training and validation folder. The training folder consists of 80% images from the train folder, and the remaining 20% images are in the validation folder. Open- and closed-eye images are distributed evenly across each folder. Dataset distribution diagram for model training, testing and validation is shown in figure 5.

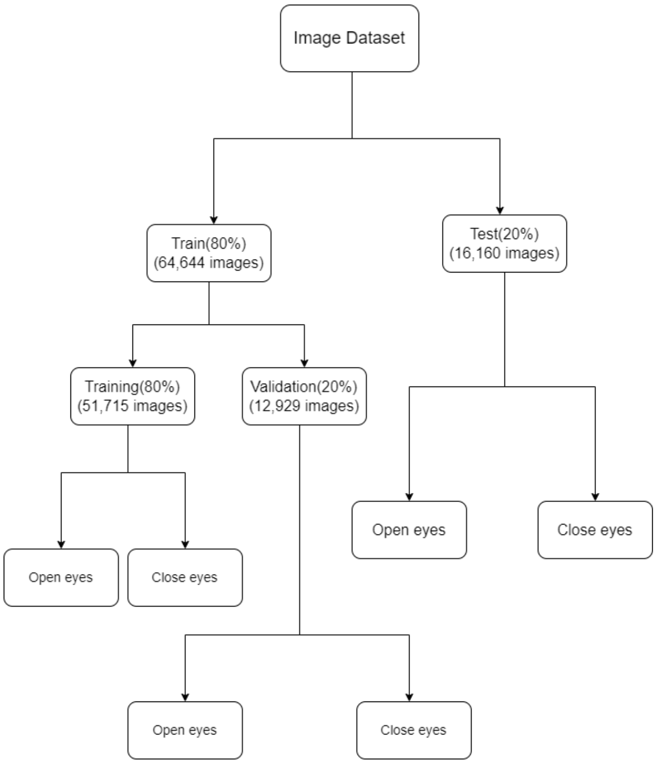

Fig. 5. Dataset distribution diagram

Images in the training dataset are rotated by 20%, shear by 20%, zoom by 20%, and the width and height of the images are manipulated by 20% and 50%, respectively, so that versatility increases in the dataset. In addition, each image is rescaled, and every pixel is divided by 255 so that all images are cropped down to a particular level. This will reduce the difficulty of the model in recognizing the face and eyes in the given dataset [11][12].

C. Model Training

Model training is a fundamental process in machine learning and deep learning where a model learns patterns and relationships within data to make predictions or perform tasks. CNN's InceptionV3 architecture is used in our proposed model to train the model. In the base model, the last layers of the InceptionV3 architecture are not included as our model is predicated on knowledge transmission and we only need knowledge from a pre-trained model [13].

The head model is the output of the base model. After flattening the head model and taking only important data from it, head_model = dropout(0.5) is set so that the model does not overtrain. At last, two neurons are added to detect eye state, i.e., open or closed state.

The proposed model gets input from the base model and output from the head model. The total number of parameters that we can play over is 21,934,050 and the trainable parameters are 131,266 which is shown in figure 6.

```
==================================================
Total params: 21,934,050
Trainable params: 131,266
Non-trainable params: 21,802,784
__________________________________________________
```

Fig. 6. Parameter extracted from model

The proposed model imports ModelCheckpoint, EarlyStopping, and ReduceLROnPlateau. ModelCheckpoint is used to save only the best model, EarlyStopping is used to stop the model from overtraining, and ReduceLROnPlateau is used to reduce the learning rate if the model accuracy is not increasing after 3 epochs. The model will stop training if the model's accuracy is not increasing rather decreasing after 7 epochs and the model's best performance is saved. Batch size is set to 128 and epochs is set to 50.

D. Alarm Triggering Mechanism

In the developed drowsiness detection system, the alarm triggering mechanism serves as a vital component to alert the driver when prolonged eye closure indicative of potential drowsiness is detected. Through the dashboard camera, a video of the driver while driving in real time is captured by the proposed model and real-time output is produced.

a) Functionality:

- Monitoring Eye Closure Duration: The system continuously monitors the driver's eyes using the CNN InceptionV3 model. It tracks the duration of eye closure in real-time during the driver's interaction with the system.

- Scoring System: An algorithm within the system calculates a 'score' based on the duration of eye closure. This 'score' acts as an accumulative measure, incrementing as the duration of eye closure extends.

- Threshold-Based Alarm: When the 'score' surpasses a predefined threshold value of 15 , the system triggers an alarm. when the driver opens the eye, then this score decreases automatically and when it comes below 15, the alarm stops.

b) Alerting the Driver:

- Alarm Activation: Upon reaching the threshold, the system activates an alarm, typically in the form of a beep or a distinctive sound.

- Driver Notification: The purpose of the alarm is to alert the driver, bringing attention to the potential onset of drowsiness due to prolonged eye closure. The alert aims to prompt the driver to take corrective actions, such as taking a break or adjusting their focus and attentiveness while driving.

Flowchart of the proposed model is shown in figure 7.

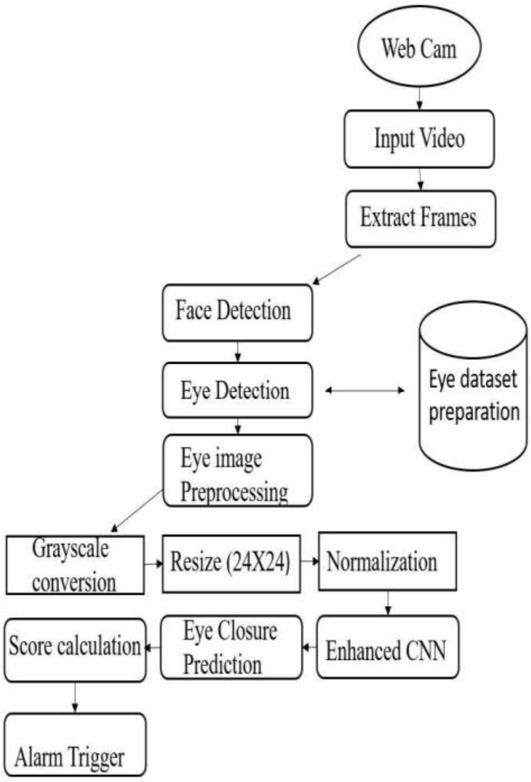

Fig. 7. Flowchart of driver drowsiness detection proposed model

III. MODEL PERFORMANCE ANALYSIS

The model's accuracy and loss in the training, validation, and test datasets is used to determine the model's performance.

A. Accuracy Evaluation:

- Training Accuracy (88.694%): The model achieved a high training accuracy of approximately 88.69%. This signifies that during training, the model correctly predicted the class labels for around 86.69% of the images in the training dataset.

- Validation Accuracy (81.564%): The validation accuracy stands at about 81.564%. Even though it is somewhat less accurate than the training data, it nevertheless shows that the model can generalize to newly discovered validation data..

- Test Accuracy (83.975%): The test accuracy, at approximately 83.975%, reflects the model's performance on a separate test dataset not used during training. It showcases the model's generalization ability to new, unseen data.

B. Loss Evaluation:

- Training Loss (1.063): The training loss value of 1.063 indicates the average loss calculated during model training. Lower values denote better convergence and alignment between predicted and true values.
- Validation Loss (1.181): The validation loss of 1.181 on the validation dataset represents the model's performance on new, unseen validation data. It's higher than the training loss, suggesting some potential overfitting or difficulty generalizing.
- Test Loss (1.002): The test loss value of 1.002 reflects the model's loss on the separate test dataset. A lower test loss compared to the training and validation losses is positive, indicating good performance on new data.

C. Training Dynamics:

- Epochs and Learning Rate Adjustment: The model trained for a total of 12 epochs, with a learning rate reduction after 7 epochs since the loss value didn’t improve. This adjustment might suggest that the model initially made considerable progress, but its improvement slowed down, leading to a decreased learning rate for fine-tuning.
- Early Stopping: The model stopped training after 12 epochs, potentially indicating that further training did not significantly enhance its performance.

D. Batch Size and Dataset Size:

- Batch Size (128): A larger batch size of 128 was used for training, which generally provides stability in learning and computational efficiency compared to smaller batch sizes.
- Dataset Size (64,644 images): The training dataset consisting of 61,325 images is significant, allowing the model to learn diverse patterns present in the data.

IV. RESULT AND DISCUSSION

A. Result

Any deep learning model may be assessed using the performance matrices. Here, in our research, we have evaluated the accuracy and performance of our CNNs models utilizing ideas like accuracy measure and loss measure.. Model accuracy and loss is shown in figure 8.

```
In [65]: training_loss,training_accuracy = model.evaluate_generator(train_data)
         validation_loss,validation_accuracy= model.evaluate_generator(validation_data)
         test_loss,test_accuracy = model.evaluate_generator(test_data)

In [66]: print(training_accuracy)
         print(training_loss)
         print(validation_accuracy)
         print(validation_loss)
         print(test_accuracy)
         print(test_loss)

         0.88694323
         1.6034534231132453
         0.81564256
         1.1816357823410324
         0.8397534
         1.002870395151342
```

Fig. 8. Model accuracy and loss

The Output from the model is shown in figure 9 and figure 10. Figure 9 shows the score is 0 when eyes are open thus no alarm beeping. Figure 10 shows the score is 22 when eyes are closed thus alarm beeps to alert the diver.

B. Discussion

The model's consistency across accuracy metrics, along with its low-test loss, indicates its ability to effectively detect drowsiness indicators, showcasing promise for real-world deployment within automotive safety systems.

The slight discrepancy between training and validation accuracies suggests some degree of generalization, although the gap might signify slight overfitting or limitations in capturing certain nuances present in the validation dataset.

The model's convergence and stability throughout the 50 epochs provide confidence in its learning capabilities. However, further optimization and fine-tuning may be considered to address potential overfitting and enhance generalization to improve the model's performance across all datasets.

The overall performance of the model positions it as a viable candidate for integration into real-world driver safety systems, contributing significantly to mitigating the risks associated with drowsy driving. Continued refinement and exploration of diverse strategies may further elevate the model's accuracy and robustness, ensuring its effectiveness in practical deployment scenarios.

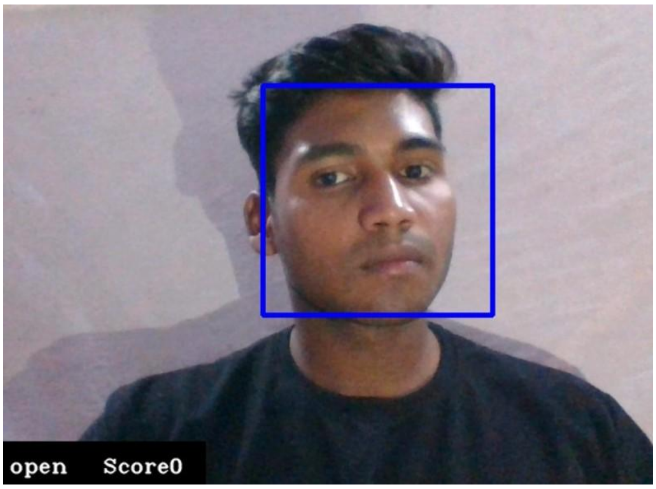

Fig. 9. Score 0 when eyes are open

Fig. 10. Score 22 when eyes are closed

V. CONCLUSION AND FUTURE DIRECTION

A. *Conclusion*

The research has successfully developed and evaluated a driver drowsiness detection model leveraging InceptionV3 architecture by using Convolutional neural network and transfer learning with promising performance metrics. The system demonstrates 84% accuracy in identifying drowsiness symptoms in driver, contributing significantly to automotive safety systems.

a) Key Findings:

- Notable accuracy achieved in detecting drowsiness indicators based on eye closure duration.
- Successful deployment considerations for real-time implementation within vehicles.

Proposed model can be deployed in vehicle to reduce accidents due to tiredness and lack of attention of driver. Drivers real time video will be analysed frame by frame by the model and real time output i.e. alarming sound will be produced to alert the driver if driver is being sleepy.

B. *Future Direction*

- Continuous Learning Mechanism: Develop mechanisms for the model to adapt and learn from new data in real-time, ensuring ongoing improvement and adaptability.
- Personalized Drowsiness Detection: Investigate methods to personalize the detection system based on individual variations in drowsiness indicators, improving accuracy for diverse drivers.
- Steering pattern monitoring: primarily makes use of the electric power steering system's steering input. This kind of monitoring a driver is only effective when the driver actively steers the vehicle.

ACKNOWLEDGMENT

We extend our heartfelt gratitude to Chandigarh University for providing us with the platform to undertake our project on the subject of "Driver Drowsiness detection using Convolutional Neural Network". We also wish to express our appreciation to Dr. Navneet Kaur, our Head of Department, for granting us the opportunity and resources to delve into this research endeavor. We are deeply thankful to Prof. Prabhjot Kaur from the Department of Computer Science and Engineering for her invaluable guidance and unwavering support throughout our research journey. Her motivation and insightful suggestions were instrumental in making this research a reality. Furthermore, we acknowledge that this research paper would have been a far more challenging undertaking without the cooperation, suggestions, and assistance from our dedicated circle of friends and the unwavering support of our families.

REFERENCES

[1] Rivelli, E. (no date) *Drowsy driving 2021 Facts & Statistics*, *Bankrate*. Available at: https://www.bankrate.com/insurance/car/drowsy-driving-statistics/ (Accessed: 17 November 2023).

[2] Shekari Soleimanloo, S. *et al.* (2019) 'Eye-blink parameters detect on-road track-driving impairment following severe sleep deprivation', *Journal of Clinical Sleep Medicine*, 15(09), pp. 1271–1284. doi:10.5664/jcsm.7918.

[3] Simons-Morton, B.G. *et al*. (2014) 'Keep your eyes on the road: Young driver crash risk increases according to duration of distraction', *Journal of Adolescent Health*, 54(5). doi:10.1016/j.jadohealth.2013.11.021.

[4] Baek, J.W. *et al.* (2018) 'Real-time drowsiness detection algorithm for driver State Monitoring Systems', *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)* [Preprint]. doi:10.1109/icufn.2018.8436988.

[5] Sharma, N., Jain, V. and Mishra, A. (2018) 'An analysis of convolutional neural networks for Image Classification', *Procedia Computer Science*, 132, pp. 377–384. doi:10.1016/j.procs.2018.05.198.

[6] *Introduction to convolution neural network* (2023) *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/introduction-convolution-neural-network/ (Accessed: 17 November 2023).

[7] *Transfer learning with Convolutional Neural Networks* (no date) *Online Tutorials, Courses, and eBooks Library*. Available at: https://www.tutorialspoint.com/transfer-learning-with-convolutional-neural-networks (Accessed: 18 November 2023).

[8] Sharma, P. (2023) *Understanding transfer learning for deep learning*, *Analytics Vidhya*. Available at: https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/ (Accessed: 17 November 2023).

[9] Szegedy, C. *et al*. (2015) *Rethinking the inception architecture for computer vision*, *arXiv.org*. Available at: https://arxiv.org/abs/1512.00567 (Accessed: 17 November 2023).

[10] *MRL Eye Dataset* (no date) *MRL*. Available at: http://mrl.cs.vsb.cz/eyedataset (Accessed: 18 November 2023).

[11] Traore, B.B., Kamsu-Foguem, B. and Tangara, F. (2018) 'Deep Convolution Neural Network for Image Recognition', *Ecological Informatics*, 48, pp. 257–268. doi:10.1016/j.ecoinf.2018.10.002.

[12] Traore, B.B., Kamsu-Foguem, B. and Tangara, F. (2018a) 'Deep Convolution Neural Network for Image Recognition', *Ecological Informatics*, 48, pp. 257–268. doi:10.1016/j.ecoinf.2018.10.002.

[13] Xin, M. and Wang, Y. (2019) 'Research on image classification model based on Deep Convolution Neural Network', *EURASIP Journal on Image and Video Processing*, 2019(1). doi:10.1186/s13640-019-0417-8.