Driver's Drowsiness Detection Using Deep Learning Approach

Jalpa Kaila
Department of Computer Science and Engineering
University of Nevada, Reno
jalpakaila88@nevada.unr.edu

May 2024

Abstract

Driver's drowsiness is one of the leading causes of traffic accidents and significantly affecting overall road safety. Providing early warnings to drowsy drivers could prevent numerous traffic accidents. Drowsiness detection systems observe the driver's state and trigger an alert upon detecting signs of drowsiness. To contribute to the current works of drowsiness detection systems, this work aims to study a Driver's Drowsiness Detection using Deep Learning approach. This paper proposes a system for driver drowsiness detection, in which the architecture detects sleepiness of driver. The proposed architecture consists of two deep learning models: VGG-16 and InceptionV3 which uses Grayscale images of drivers as input and help in detecting drowsiness. The MRL Eye dataset is used for the Experiment with VGG16 and InceptionV3 Network which contains about 4000 images and separated into two folder Closed and open Eyes. The experiment was conducted using train, validation, and test sets comprising 70-20-10% and 60-20-20% of images, respectively. Following that, a SoftMax classifier will be applied to determine the drowsiness or non-drowsiness behavior of a driver.

1 Introduction

Drowsiness is a big problem on the road because when drivers start dozing off, they might not realize it. This can lead to accidents because they're not fully aware. Feeling sleepy while driving can happen if you haven't slept enough, feel tired, or are dealing with stress or other mental stuff. Thousands of people die every year resulting from vehicle accidents due to drowsy driving [1][2]. It was also reported that the main causes for most of these accidents were drowsiness, speeding and changing lanes suddenly, inattentiveness, and crossing red lights. In order to reduce such an accidents and enhance the safety of the driver and the passengers, driver drowsiness detection systems have been worked on and developed by various researchers all across the world. So, various techniques were proposed to build a Driver Drowsiness Detection (DDD) system that detects the driver's drowsiness and triggers an alarm when drowsiness is detected.

The Driver's drowsiness detection system categorized into three types based on the features they utilize: vehicular, visual, and non-visual features. The Vehicular type of drowsiness detection systems rely on data collected from the vehicle itself, such as steering wheel movements, vehicle speed, lane deviation, or other vehicle dynamics. The Visual type of drowsiness detection systems utilizes how the driver's face and eyes are behaving. The camera or sensors inside the vehicle are used to capture this movement. This system analyzes facial expressions, eye movements, and blink patterns to assess drowsiness levels. For instance, if the system detects frequent eye closures or drooping eyelids, it may indicate that the driver is getting sleepy. The non-visual type of drowsiness detection systems focus on detecting drowsiness using physiological signals or behavioral patterns that are not visually observable. Examples of non-visual features include heart rate variability, electroencephalogram (EEG) signals, or even voice analysis.

In recent years, advancements in deep learning techniques have paved the way for more accurate and efficient drowsiness detection systems. Among these techniques, convolutional neural networks (CNNs) have demonstrated remarkable capabilities in analyzing and extracting meaningful features from image data. This study proposed a system for driver's drowsiness detection, in which model architecture detects sleepiness of driver. It utilizes two widely acclaimed CNN architectures: VGG16 and InceptionV3 to detect the accuracy of the classification.

2 Related Work

Over the past decade, researchers worldwide have suggested numerous methods to solve the drowsiness detection problem. One of these techniques depends on visual-based measures. Drowsy people exhibit several visual behaviors that can be detected through changes in their facial features such as eye, mouth and head [3][4]. To detect if a driver is sleepy or awake, visual behaviors are crucial features because they are quick to detect and provide clear signs through the driver's facial expressions [4][5]. Furthermore, the utilization of these visual measures is widely accepted as an effective method for detecting drowsiness, mainly because they are non-intrusive [6]. Many Visual features such as eye blinking, yawning and head movement can be detected from driver's face image [6][7].

The system that detect driver's drowsiness based on eyes and mouth focus on pupil movement and yawning parameters that measured based on information of eye and mouth [8][9]. Driver's pupil movement can be detected by movement and state of eyes. Driver's yawning parameter can be detected by width of mouth opening and lips position etc [9]. This system works as follows: Initially the face is located through Viola-Jones face detection method to ensure the presence of driver in video frame. Then, a mouth window is extracted from the face region and spatial fuzzy c-means clustering is applied to search and locate lips in the window. At the same time, the upper portion of the facial window identifies pupils based on radii, inter-pupil distance, and angle. The observed data from the eyes and mouth are further passed to Support Vector Machines (SVM), which classify the actual state of the driver. One more important feature is Detecting a Eye-blink. This method detects drowsiness by comparing the eye blinking frequency of a driver with the normal rate (around ten blinks per minute), noting decreased blinking when drowsy [10].

An effective real-time algorithm for detecting eye blinking using facial landmark detectors is proposed by Soukupova and Cech [11]. The study introduced a scalar quantity known as Eye Aspect Ratio (EAR), which is based on the ratio of the height to the width of the eye. The EAR used to measure, how open the eyes are in every frame, and it identifies a blink when there's a sudden decrease in the EAR value. This experiment uses SVM classifier to detect eye blinks as a pattern of EAR values in a short temporal window. Detection based on head movements is another prevalent feature in visual-based Driver's Drowsiness Detection systems. As drivers become drowsy, their head begins to sway and the vehicle may wander away from the center of the lane. In [10], another characteristic is noted, which indicates that the head-nodding frequency increases as the drivers nod their heads more frequently trying to avoid sleeping.

Additionally, [12] suggests that various head angles relative to the forehead can serve as another distinguishing characteristic. The sensor called Micro-Nod detection system which detects micro-sleep by analyzing the head position in real-time. The change in the driver's head position is tracked on x, y, and z coordinates. Then handled by ML algorithms such as Convolution Neural Network, Support Vector Machine, and Hidden Markov model [10][13]. The Detection of Driver's Drowsiness using a Machine Learning approach involves the detection of the face region from a video dataset [10]. The eye region is then detected and extracted as a region of interest using a facial landmark detector. Subsequently, the EAR value of each frame is calculated and three different classifiers, namely, linear support vector machine, random forest, and sequential neural network, are employed to improve the detection accuracy.

The study in [14] utilized a Deep Neural Network for Drowsiness Detection, employing features extracted by a convolutional neural network. A softmax layer classified drivers as drowsy or non-drowsy, achieving 78% accuracy across diverse datasets. Another approach, discussed in [15], involved Convolutional Neural Networks for drowsiness detection as an object detection task. This method localized open and closed eyes from a driver's video stream, using the MobileNet CNN architecture. Training the MobileNet Network involved a customized dataset of around 6000 images, labeled with face, open eye, and closed eye objects. This methodology, besides maintaining accuracy, proved computationally efficient and cost-effective.

The ensemble approach for driver's drowsiness detection, detailed in [16], combines four deep learning models: AlexNet, VGG-FaceNet, FlowImageNet, and ResNet. These models analyze RGB videos of drivers, considering various features like hand gestures, facial expressions, behavioral traits, and head movements. AlexNet handles diverse environmental conditions, while VGG-FaceNet extracts facial characteristics such as gender and ethnicity. FlowImageNet focuses on behavioral features and head gestures, while ResNet identifies hand gestures. The ensemble algorithm integrates their outputs through a SoftMax classifier, producing a final classification of drowsy or non-drowsy. This approach offers comprehensive detection by leveraging multiple models and diverse features, enhancing accuracy in real-world scenarios.

3 Methodology

The project aims to develop a Driver's Drowsiness Detection system using two Deep Learning architecture VGG16 and InceptionV3. From the study of [10] which involves the EAR method for feature extraction as shown in Figure 1 and Study [14], [16] which involves the Deep learning based feature learning model, involves detecting and extracting the eye region from images using a facial landmarks detector. The EAR values are then calculated and analyzed. It uses Sequential Neural Network for classification, with a SoftMax classifier used to determine drowsiness behavior. The performance obtained in [10], [14] and [16] with Existing method and different data set is satisfactory for Experiment on Driver's Drowsiness Dataset.

Convolutional Neural Networks (CNN), also known as ConvNets, are powerful deep learning algorithms widely employed in image processing tasks. They learn unique weights and biases to distinguish between different objects within input images. Operating on pixel value matrices representing image dimensions, CNNs consist of layers including convolutional, activation, and pooling layers. Convolutional layers employ kernels to scan image regions, capturing diverse features. Activation layers utilize decision functions to analyze image nonlinearity, while pooling layers downsample images. Sequentially, CNNs incorporate fully connected (FC) layers after convolutional, activation, and pooling layers, producing N-dimensional vectors corresponding to N output classes. This architecture enables CNNs to effectively differentiate between images, making them indispensable in various image recognition applications.

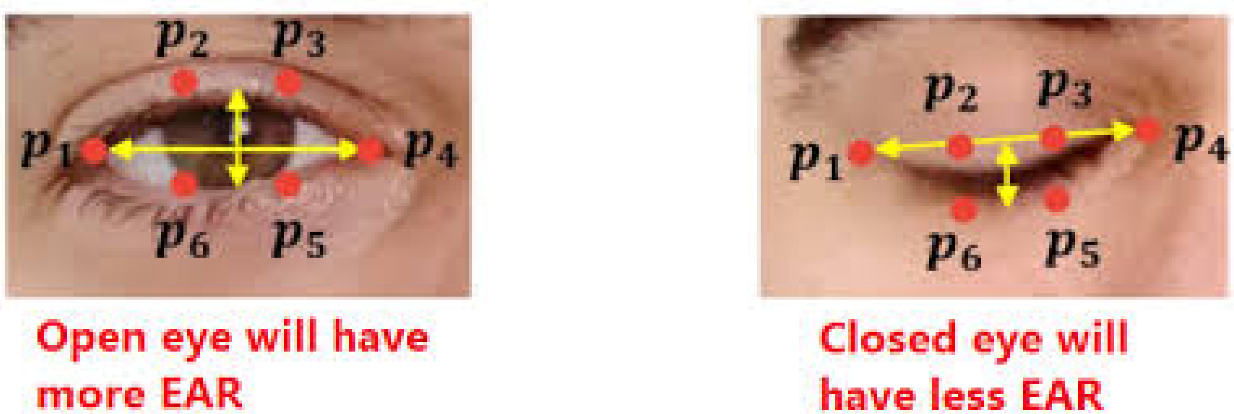

Figure 1: Eye Aspect Ratio

3.1 Architecture and Dataset

The process of developing a Driver's Drowsiness Detection system using Deep Learning involves several steps, including data collection, categorization, model selection, training, and evaluation. Flow chart shown in figure 2 involves each process. The MRL Eye Dataset is a comprehensive collection of eye images is used for this study. It features annotations for various properties, including subject ID, image ID, gender, presence of glasses, eye state (open or closed), reflections (none, small, or big), lighting conditions (bad or good), and sensor ID. The dataset encompasses a wide range of conditions, including different genders, presence of glasses, and variations in lighting conditions and reflections[17].

3.2 VGG16

VGG16 is a convolutional neural network (CNN) architecture known for its simplicity and effectiveness in image classification tasks. It consists of 16 layers, including convolutional layers with small 3x3 filters, followed by max-pooling layers. The network progressively learns abstract features from input images through multiple convolutional layers, capturing complex patterns and details. VGG16's deep architecture enables it to extract hierarchical representations of visual features, facilitating accurate classification. However, it

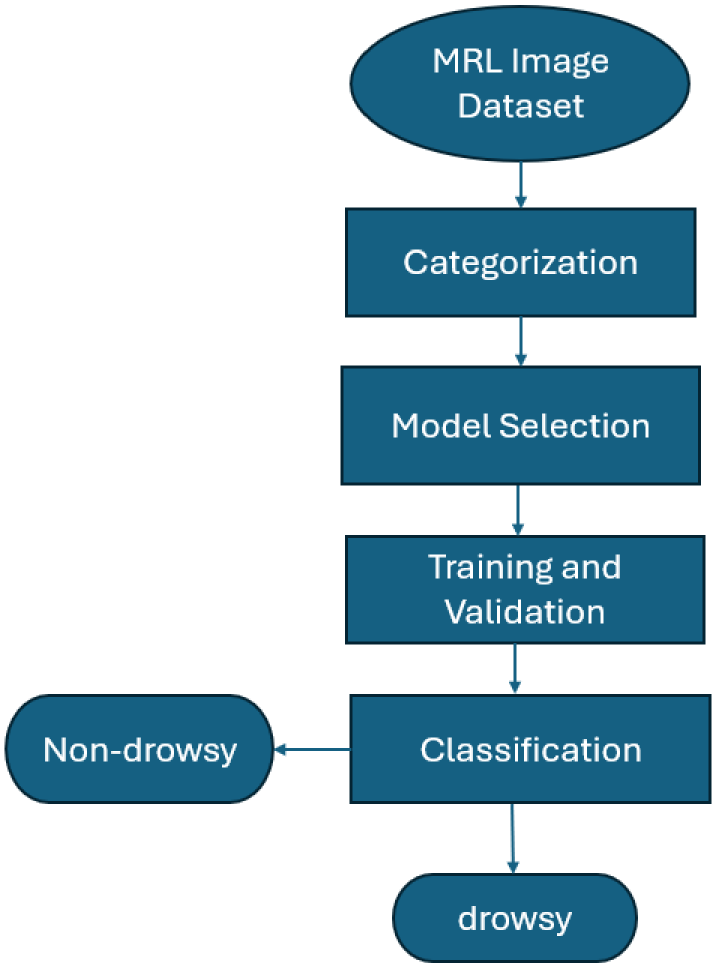

Figure 2: Process Flow Diagram

suffers from the limitation of high computational cost because of the application of approximately 140 million parameters.

3.3 InceptionV3

InceptionV3 is a convolutional neural network (CNN) architecture developed for image classification tasks. It incorporates a novel inception module, which enables efficient feature extraction through parallel operations with varying filter sizes. This design facilitates the network's ability to capture both local and global features effectively, enhancing its performance in image recognition tasks. InceptionV3 introduces factorized convolutions, which reduce computational complexity while preserving representational capacity. Additionally, the architecture employs techniques such as batch normalization and auxiliary classifiers to aid in training and improve gradient flow, resulting in faster convergence and better generalization. InceptionV3 is widely adopted choice for tasks requiring accurate and efficient feature extraction from visual data, including image classification, object detection, and semantic segmentation.

3.4 Experimental Setup

MRL dataset with 4000 images categorized in two folders Closed and Open Eyes. In Experiment VGG16 and InceptionV3 pre-trained model is used and trained it on 4000 images of MRL dataset (with Closed and open Eyes). The experiment was conducted using train, validation, and test sets comprising 70-20-10% and 60-20-20% of images, respectively. TensorFlow's Keras API is utilized for constructing a convolutional neural network (CNN) using the pre-trained VGG16 and InceptionV3 model for drowsiness detection tasks. The batch size used here for both models is 8, which means during each iteration of training, the model will process 8 images simultaneously. It helps in optimizing memory usage and computational efficiency. The number of epochs used in this experiment for training both models is 5. An epoch is one complete pass through the entire training dataset. Training for multiple epochs allows the model to learn from the data iteratively and improve its performance over time. Both the model is augmented with additional layers, including dense and dropout layers, to customize it for the specific task at hand. Following that, a SoftMax classifier is applied to determine the drowsiness or non-drowsiness behavior of a driver. The selected model is trained on the training data using appropriate loss functions and Monitor the model's performance on the validation set. The trained models Evaluated on the test set to assess their performance in drowsiness detection. Calculate metric such as accuracy and loss are used to measure the model's effectiveness. The hypotheses of this study are as follows:

- **Hypothesis 1:** Evaluate the performance of VGG16 in detecting drowsiness for MRL Dataset.
- **Hypothesis 2:** Evaluate the performance of InceptionV3 in detecting drowsiness for MRL Dataset.

4 Results and Discussion

This Section represents the results of VGG16 and InceptionV3 using MRL dataset which divided into train, validation, and test sets comprising 70-20-10% and 60-20-20% of images, respectively. Figure 3 and 4 represents the samples of 5 predicted images from test set using VGG16 and we can see that in Figure 4, first image is actually open eyes is predicted as closed eyes, thus it is incorrectly classified. The remaining images are correctly classified. Figure 5 and 6 represents the samples of 5 predicted images from test set using InceptionV3 model. We can see that, all the images are correctly classified. As InceptionV3 is more accurate it gives correct prediction with both cases on Datasets.

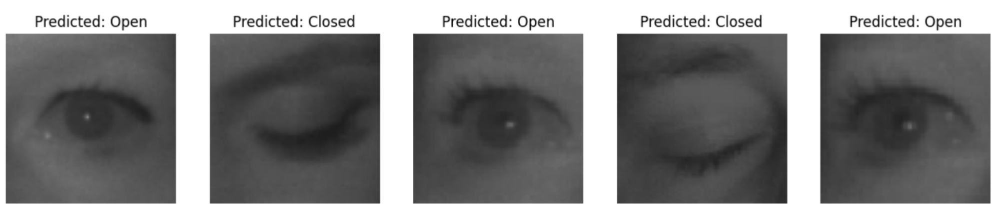

Figure 3: Predicated 5 images using VGG16 (70% 20% 10%) Dataset

Figure 4: Predicated 5 images using VGG16 (60% 20% 20%) Dataset

Figure 5: Predicated 5 images using InceptionV3 (70% 20% 10%) Dataset

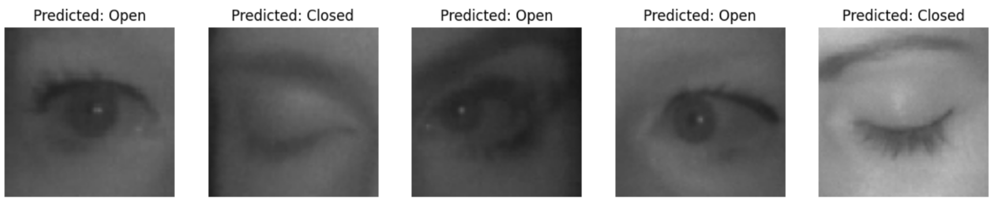

Figure 6: Predicated 5 images using InceptionV3 (60% 20% 20%) Dataset

Table 1 represents the Accuracy and Loss Values of VGG16 and InceptionV3 using MRL dataset in both 70-20-10% and 60-20-20% of images. The accuracy for both the model in case of 70-20-10% Dataset is more compared to 60-20-20%. Though Inception Model gives better accuracy than VGG16 in both the cases. With these experimental results, we proved both hypotheses.

Model	Accuracy			Loss		
	Train	Validation	Test	Train	Validation	Test
VGG16 (70%20%10%)	96.49%	95.37%	95.74%	3.51%	4.63%	4.26%
VGG16 (60%20%20%)	96.54%	78.37%	93.37%	3.46%	21.63%	6.63%
InceptionV3 (70%20%10%)	99.32%	99.03%	99.01%	0.68%	0.97%	0.99%
InceptionV3 (60%20%20%)	98.20%	98.00%	98.75%	1.80%	2.01%	1.25%

Table 1: Analysis of Accuracy and Loss

5 Conclusion and Future work

This study demonstrates the efficacy of employing the InceptionV3 model over VGG16 for drowsiness detection using the MRL dataset. Based on the results obtained from testing the dataset under various illumination, reflection, and view-point conditions of the MRL dataset, we conclude that the proposed methodology employing the InceptionV3 model yields better results than the VGG16 model for drowsiness detection. Both hypotheses are validated in our experiment, and the accuracy and loss of both models are effectively analyzed. The InceptionV3 model outperformed VGG16 across various illumination, reflection, and different view-point conditions, showcasing its superior capabilities in accurately detecting drowsiness.

One major improvement that could be made in the future is to experiment with different datasets and models to improve accuracy and generalization. Additionally, making the system applicable in the real world could involve the incorporation of alarm beeps to alert drivers when drowsy behavior is detected, thus enhancing road safety.

References

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

2. National Safety Council (2020) Drivers are falling asleep behind the wheel https://www.nsc.org/road-safety/safety-topics/fatigueddriving

3. G. Sikander and S. Anwar, "Driver Fatigue Detection Systems: A Review," In IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 6, pp. 2339-2352, June 2019.

4. B. G. Pratama, I. Ardiyanto and T. B. Adji, "A Review on Driver Drowsiness Based on Image, Bio-Signal, and Driver Behavior," In 2017 3rd International Conference on Science and Technology - Computer (ICST), Yogyakarta, 2017, pp. 70-75, doi: 10.1109/ICSTC.2017.8011855.

5. S. Kaplan, M. A. Guvensan, A. G. Yavuz, and Y. Karalurt, "Driver Behavior Analysis for Safe Driving: A Survey," In IEEE Transactions on Intelligent Transportation Systems, 2015, pp. 3017-3032.

6. Q. Ji and X. Yang, "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance," Real-Time Imaging, vol. 8, no. 4, pp. 357–377, 2002.

7. A. Mittal, K. Kumar, S. Dhamija, and M. Kaur, "Head movement-based driver drowsiness detection: A review of state-of-art techniques," In 2016 IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, India, 2016, pp. 903-908.

8. M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas and A. Mahmood, "A Survey on State-of-the-Art Drowsiness Detection Techniques," IEEE Access, vol. 7, pp. 61904-61919, 2019.

9. T. Azim, M. A. Jaffar and A. M. Mirza, "Automatic Fatigue Detection of Drivers through Pupil Detection and Yawning Analysis," 2009 Fourth International Conference on Innovative Computing, Information and Control (ICICIC), Kaohsiung, 2009, pp. 441-445, doi: 10.1109/ICICIC.2009.119.

10. A. Al Redhaei, Y. Albadawi, S. Mohamed and A. Alnoman, "Realtime Driver Drowsiness Detection Using Machine Learning," 2022 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 2022, pp. 1-6, doi: 10.1109/ASET53988.2022.-9734801.

11. T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks," in 21st Computer Vision Winter Workshop. Slovenia, 2016, pp 1-8.

12. S. Arun, S. Kenneth and M. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," MDPI. Sensors, vol. 12, pp. 16937-16953. doi:10.3390/s121216937

13. M. Ngxande, J. Tapamo, and M. Burke, "Driver drowsiness detection using Behavioral measures and machine learning techniques: A review of state-of-art techniques," In 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), Bloemfontein, South Africa, 2017, pp. 156–161.

14. K. Dwivedi, K. Biswaranjan, and A. Sethi, "Drowsy driver detection using representation learning," Souvenir 2014 IEEE Int. Adv. Comput. Conf. IACC 2014, pp. 995–999, 2014.

15. Shakeel MF, Bajwa NA, Anwaar AM, Sohail A, Khan A (2019) Detecting driver drowsiness in real time through deep learning-based object detection. In: International work-conference on artificial neural networks, Springer, Cham pp 283–296

16. M. Dua, R. Singla, S. Raj, and A. Jangra, "Deep CNN models-based ensemble approach to driver drowsiness detection," Neural Computing and Applications, vol. 33, pp. 3155-3168, 2021.

17. http://mrl.cs.vsb.cz/eyedataset