

Driver Drowsiness Detection Using MobileNet and Eye Images

Posted on October 9, 2026

View this essay online:

<https://academic-master.com/driver-drowsiness-detection-mobilenet-eye-images/>

Abstract

Most car accidents may be traced back to a sleepy driver, regardless of where you live. Accidents on the road are unfortunately commonplace because many drivers are either fatigued or haven't gotten enough rest to keep their eyes open and focus. The greatest approach to prevent drowsy driving accidents is to alert the motorist ahead of time. Drowsiness can be detected in a variety of ways. In this research, we propose a deep learning-based method for identifying sleepy drivers. Here, we employ a type of deep learning known as convolutional neural networks based transfer learning (MobileNet). To make our sleepiness predictions, we used a dataset based on information from the eyes. An automatic drowsy driver detection system has been created with the help of non-invasive machine vision principles. For this purpose, a tiny monochrome security camera is aimed at the driver's face and keeps tabs on his or her eyes. A warning signal would be given to the driver in this situation if fatigue was detected. Information about where to look for eyes and how to tell if they're open or closed is included in this report. One of the main goals of the research was to create a novel algorithm that does not exist in any previously published papers. Finding the boundaries of the face using data from the binary image helps the algorithm zero in on the likely locations of the eyes. When the whole face area has been determined, the eyes can be pinpointed by taking the horizontal average of the face area. The eyes are found by identifying the substantial intensity changes in the face, which is done with the prior knowledge that eye locations in the face display great intensity shifts.

Keywords: Drowsiness detection, Convolutional neural network, transfer learning, deep learning.

INTRODUCTION

Sleep-deprived driving is defined as the operation of a motor vehicle while suffering from cognitive impairment as a result of insufficient sleep (also known as tired driving, drowsy driving, or fatigued driving). Loss of sleep can have the same effect on the brain as being drunk, making it a leading cause of car crashes. One study from 1998 found that one in four adults admit to dozing off at the

wheel. The US DOT reports that male drivers are twice as likely as female drivers to acknowledge nodding off at the wheel. Vehicle traffic continues to rise, making automobile accidents the top cause of death across the country [8]. We all know that the driver of a vehicle ultimately bears responsibility for the integrity of the road network and the safety of all drivers on the road [9] [16] [19].

2. LITERATURE REVIEW

This study aims to categorise drivers' alert states, especially the slightly drowsy condition, using a combination of vehicle-based and non-vehicular sensing to identify factors like behaviour and physiological state; the ultimate goal being to combine these identifications into a single detection system [1]. This study introduces a novel approach to detecting drowsiness in drivers in real time by factoring in the person's unique traits. A deep cascaded convolutional neural network was developed for face area recognition to address the problem of low accuracy brought on by artificial feature extraction. Dlib is used as a set of tools to identify the frontal driver's facial landmarks in an image. To assess if a driver is sleepy, the new metric, Eyes Aspect Ratio, examines the relative placement of visual indicators inside the frame [2].

In this study, we detail the development of an Android app that can detect driver fatigue in real time. In order to determine if a certain motorist is in a video, the device uses image processing to locate their face. Face recognition, eye tracking, and adaptive thresholding are all used to determine whether or not a motorist is dozing off behind the wheel. The proposed strategy has been tested with machine learning algorithms to confirm its efficacy [3]. This research provides a method for automatically identifying whether a person's eyes are open or closed in a still snapshot of their face; this information could be useful in a variety of contexts, including human-computer interface design, emotion recognition, and monitoring of fatigue. In order to construct a model of the state of the eyes, this method takes advantage of a large number of feature sets that together define the wealth of information included in eye patches (with respect to both global and local forms and textures) [4].

The solution presented in this study to the problem of eye closure detection can be used in a variety of contexts, including but not restricted to areas such as human-computer interface development, emotion recognition, or drowsiness monitoring for drivers. This technique utilises the abilities of many feature sets to characterise the wide range of information contained inside eye patches in order to construct a model of the eye's condition (such as global and local shape and local texture information) [5].

In this article, we detail an approach to determining which EEG channels are most relevant to the investigation of alcoholism. After using Empirical Mode Decomposition (EMD) to extract amplitude and frequency modulated bandwidth information from the Intrinsic Mode Function (IMF), we use ensemble

subspace K-NN as a classifier to differentiate between heavy and social drinkers. The optimal channels can be selected using a harmony search algorithm. To determine how effective a DBHS strategy is, we employ the ensemble subspace K-Nearest Neighbor classifier [6]. This study proposes an image-based tiredness detection system that employs the Eye Aspect Ratio (EAR) technique to determine whether or not a motorist is drowsy. The eye-closure detection and analysis system was built with the help of the Pi camera, Raspberry Pi 4, and GPS module to provide real-time data. Microsleep, dim lights, and unusual eye movements are just some of the ways that this device may identify a tired driver [7].

3. EXISTING SYSTEM

This model emphasizes an existing method that was designed using some of the algorithms of machine learning [10] [15]. Here the process does not perform accurately with the algorithms that were used and detection is not up to the expected levels of accuracy.

Drawbacks of Existing System

1. Less feature compatibility
2. Improper detection

4. PROPOSED SYSTEM

In our proposed method we are performing the drowsiness detection. We are using the CNN based transfer learning model (MobileNet) for training the data. After training we are using the OpenCV library. First, the camera's live feed is recorded and analysed before being transmitted to a sleepiness prediction model. If the eyes are open or closed in a given picture, that fact is determined by a manual inspection. Think of a situation when one's eyes are closed for longer than the system allows [11] [17]. A warning siren will sound for the driver and any passengers if the system detects signs of fatigue. The proposed method's flowchart is depicted below.

4.1 Convolutional Neural Network

The convolution process forms the basis of our strategy. Step by step, we'll discuss feature detectors, which act as filters for the neural network. Additionally, we will go over feature maps, understanding the characteristics of such maps, detecting patterns, detecting at what levels, and mapping out the results. Figure 1 shows the model of CNN.

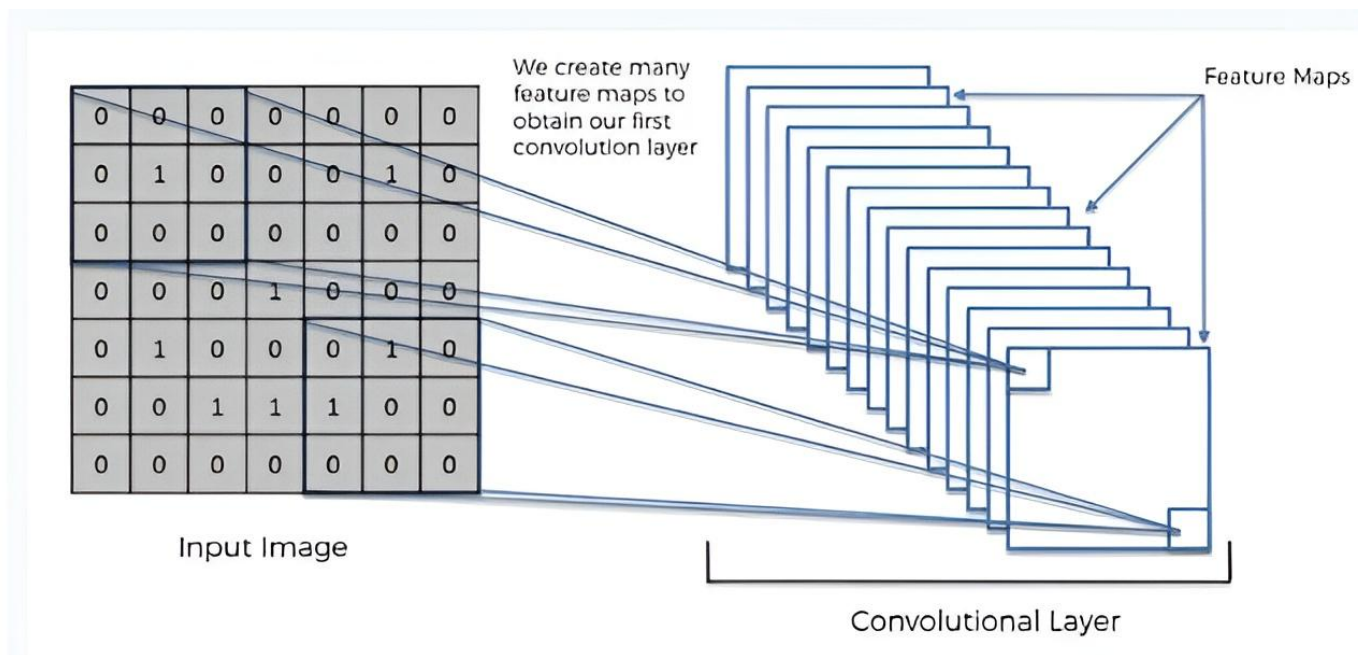


Figure 1: CNN model

4.2 ReLU Layer

The Rectified Linear Unit, also known as the ReLU, will be used in the next section of this process. In this article, we will go over ReLU layers and investigate the role of linearity in CNNs as shown in Figure 2. Not required for comprehending CNN's, but a little lesson won't hurt if you want to get better at it [12].

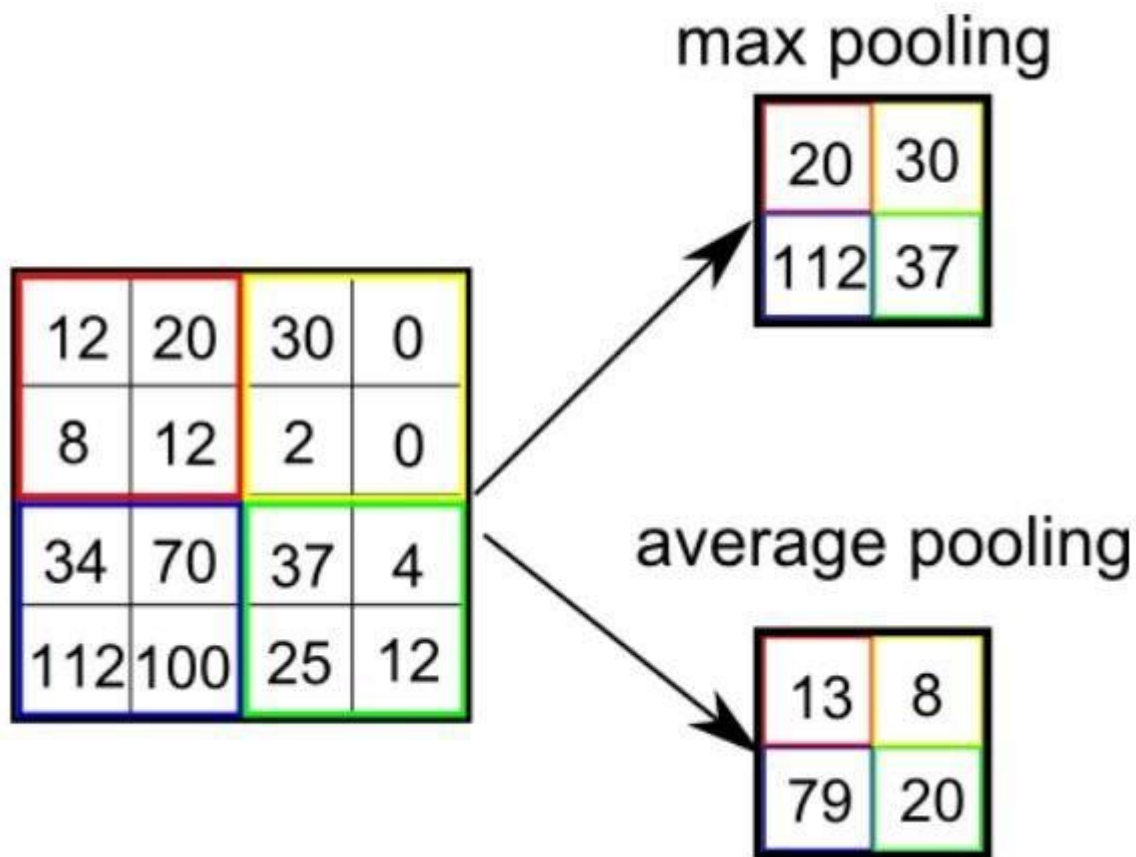


Figure 2 :ReLU Layer

4.3 Pooling Layer

This section discusses pooling and the fundamentals of its operation, focusing on max pooling. It also considers other methods, including mean pooling, sometimes described as sum pooling.

4.4 Flattening

For those unfamiliar with Convolutional Neural Networks, this section will provide a concise explanation of the flattening process and the transition from pooled to flattened layers.

4.5 Full Connection

Here, all the various topics discussed thus far will be brought together for a unified discussion. Useful for comprehending the inner workings of Convolutional Neural Networks and how the “neurons” they generate acquire the ability to categorise images as in Figure 3.

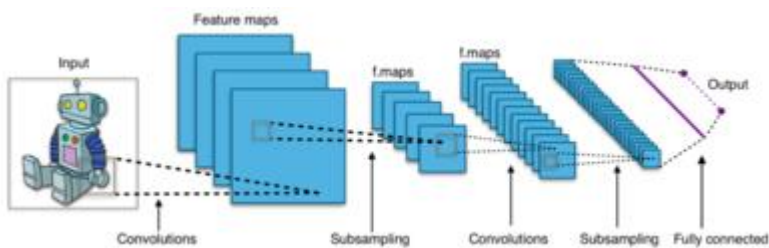
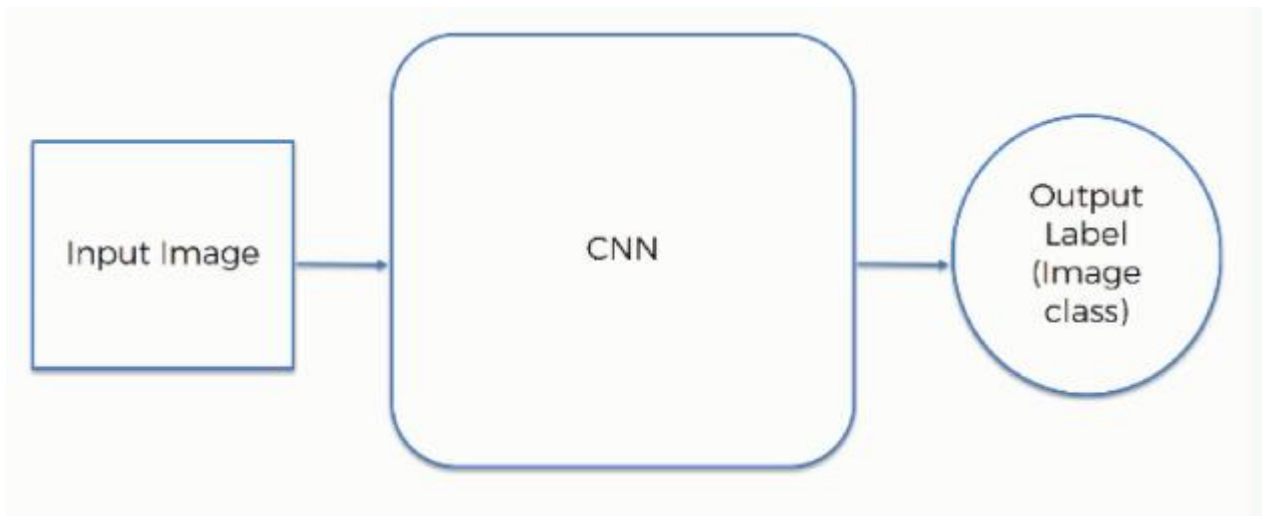


Figure 3: Full Connection

4.6 MobileNet

As a CNN architecture, MobileNet is put to use in the real world due to its efficiency and portability. To reduce model weight, MobileNets replace traditional convolutions with depthwise separable convolutions. The additional width and resolution global hyperparameters in MobileNets give model creators more flexibility in balancing competing concerns, such as latency and accuracy, against the need for speed and small size.

4.7 SYSTEM ARCHITECTURE

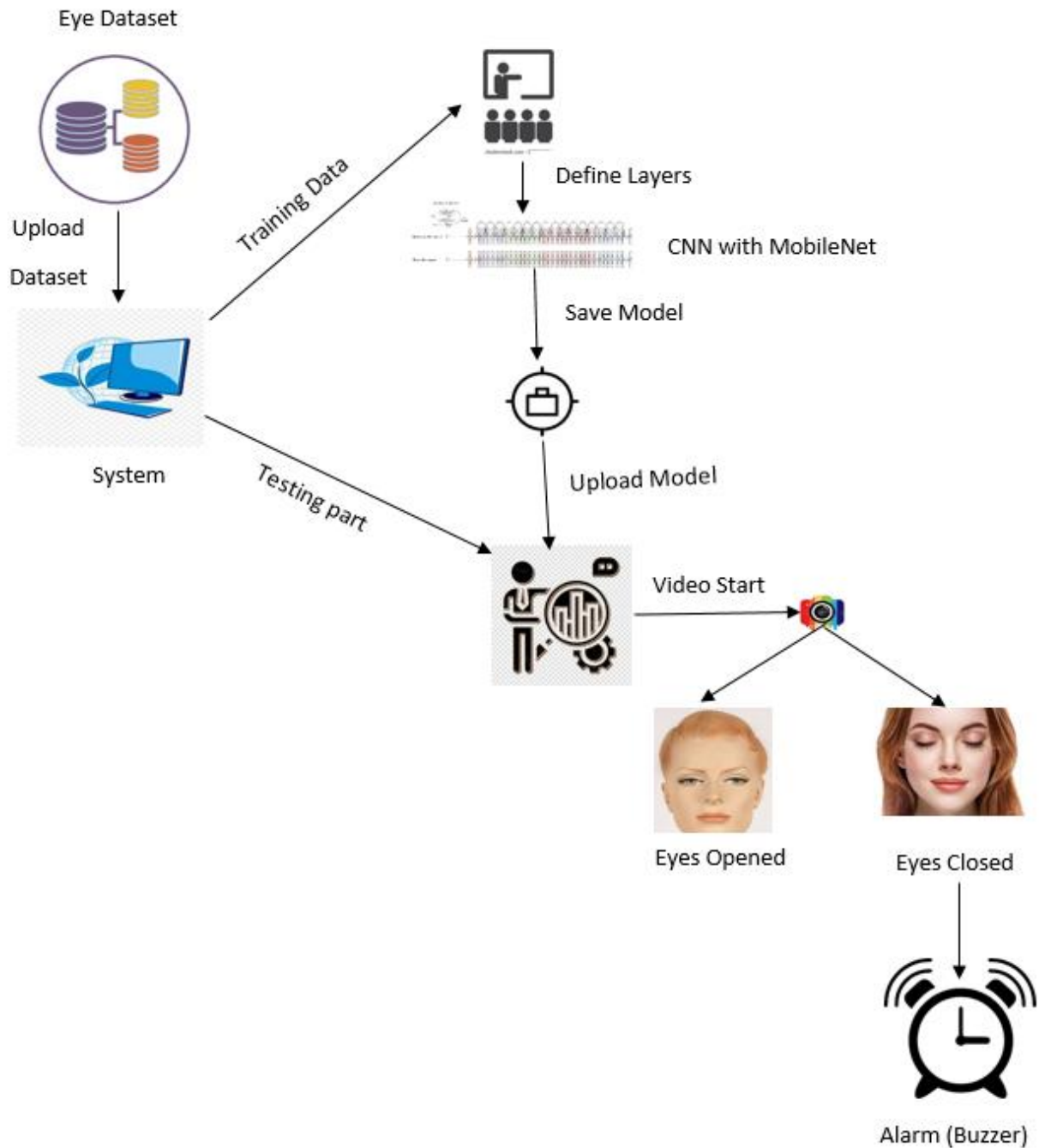


Figure 4: System Architecture

Figure 4 shows the system architecture for eye-image classification using deep learning. Features are extracted and selected from the publicly obtained dataset, then supplied to the learning algorithms to train the model. The trained model is tested with the testing data to obtain the results [13].

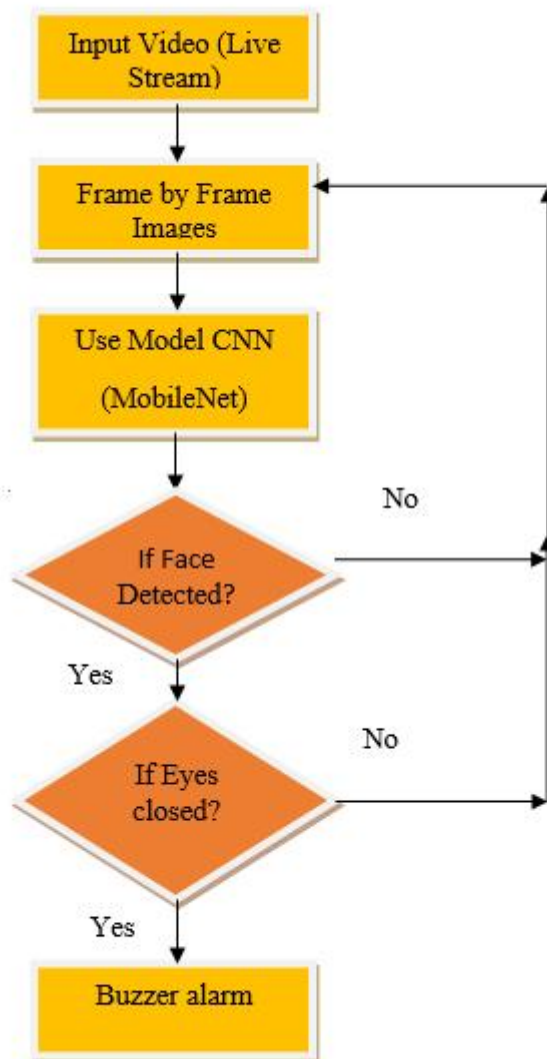


Figure 5: dataflow diagram

Figure 5 shows the pictorial representation of the architecture which the system follows to obtain the dataset, preprocess and divide the images into training and testing set.

5. LIST OF MODULES:

Create Dataset

Training

Testing

5.1 Create Dataset:

The dataset for the eye images is collected from the kaggle website (kaggle.com). The size of overall dataset is 263 MB. From there, we create two datasets, one for training and one for evaluation, with the former measuring 250MB and the latter 13MB.

5.2 Training:

To train our model, we employ CNN with the Mobile Net algorithm and a dataset that has already been preprocessed. The CNN algorithm uses the following layers: the input layer, the convolution layer, the pooling layer, the flatten layer, and the dense layer. There is an input layer where pictures are taken into consideration. Within the Convolution layer, a pixel-by-pixel image is transformed into a matrix. Here, we have a 1024-by-1024-pixel matrix (rows X columns) [14]. The numerical values will be saved in the pooling layer. For this purpose, we employ a machine learning method called Softmax to convert the numeric input to binary (supervised learning algorithm). In the Softmax layer, we'll perform the binary representation of the numerical data. Both the flatten layer and the dense layer will use a binary data format to hold the two classes that make up the entire dataset. To do this, we employ a mechanism called fit generator and save the information in .h5 files. The trained model is stored in this binary format.

5.3 Testing

By using the OpenCV technique we use the video based data through webcam. Here we will detect the faces of the individuals, it will check whether the eyes of the individuals are opened or closed. If eyes of the individuals are closed then alert will be given by playing a siren. The dataset obtained approximately 2000 train and 376 test image records of features extracted, which were then classified into 2 classes:

OPEN EYE

CLOSED EYE

5.3.1 OPEN EYE

Test data for open eye is in figure 6.





Figure 6: OPEN EYES DATASET

5.3.2 CLOSED EYE

Test data for closed-eye is in figure 7.



Figure 7: CLOSED EYE DATASET

6. PREPROCESSING

Preprocessing produces a clean dataset by removing null values, filling missing values where appropriate, removing duplicates and outliers, and discarding unwanted attributes. If the dataset contains categorical records, those variables are converted to numerical values.

In this case, we are taking a live video feed in the form of images and resizing them to a standard size.

7. BUILDING THE MODEL

SSD (Single Shot Detector) is a popular algorithm in [object detection](#).

It's generally faster than RCNN.

The "SSD head" and "backbone model" are the two main parts of an SSD. As a

feature extractor, a backbone model is often a pre-trained image categorization network

Here, we will use MobileNet SSD model to detect the objects.

Here, VGG Net is used as a backbone model to extract the features from the images.

Convolution layers (CNN) are then used for object detection in the images using the feature map generated by VGG net layer.

The model is able to detect multiple objects in any given image.

For the purpose of classification, the model uses softmax in the last layer.

Softmax takes in a vector of numbers and converts them to probabilities which are then used for image generating results.

8. PREDICTION

A live video feed is captured frame by frame as individual images. These images are preprocessed and fed into the model to predict the driver's state.

SCREENSHOT FOR DROWSINESS YAWN

Figure 8 shows the sample of drowsiness yawn followed by accuracy graph as in fig 9.

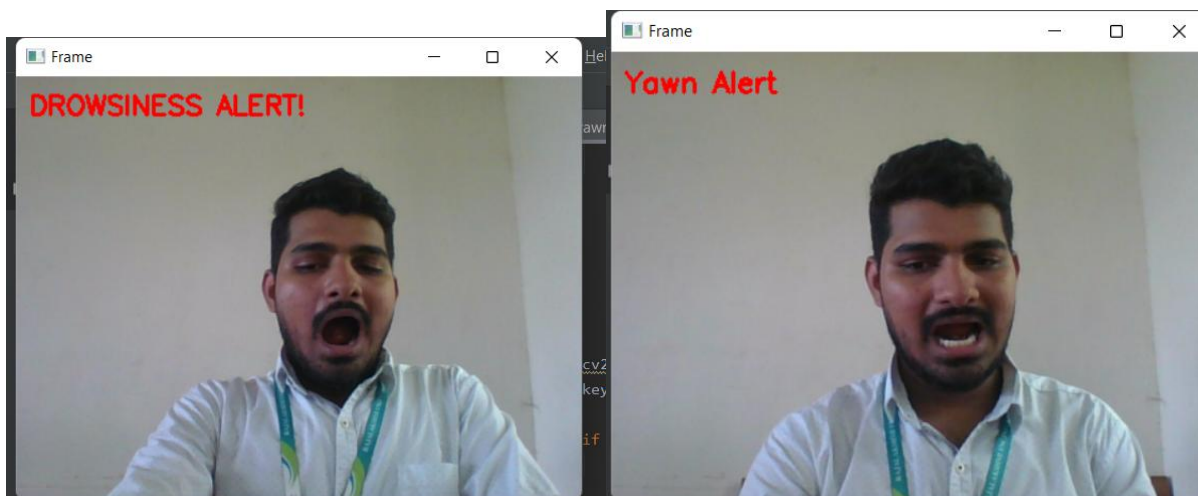


Figure 8: Screenshot for drowsiness yawn

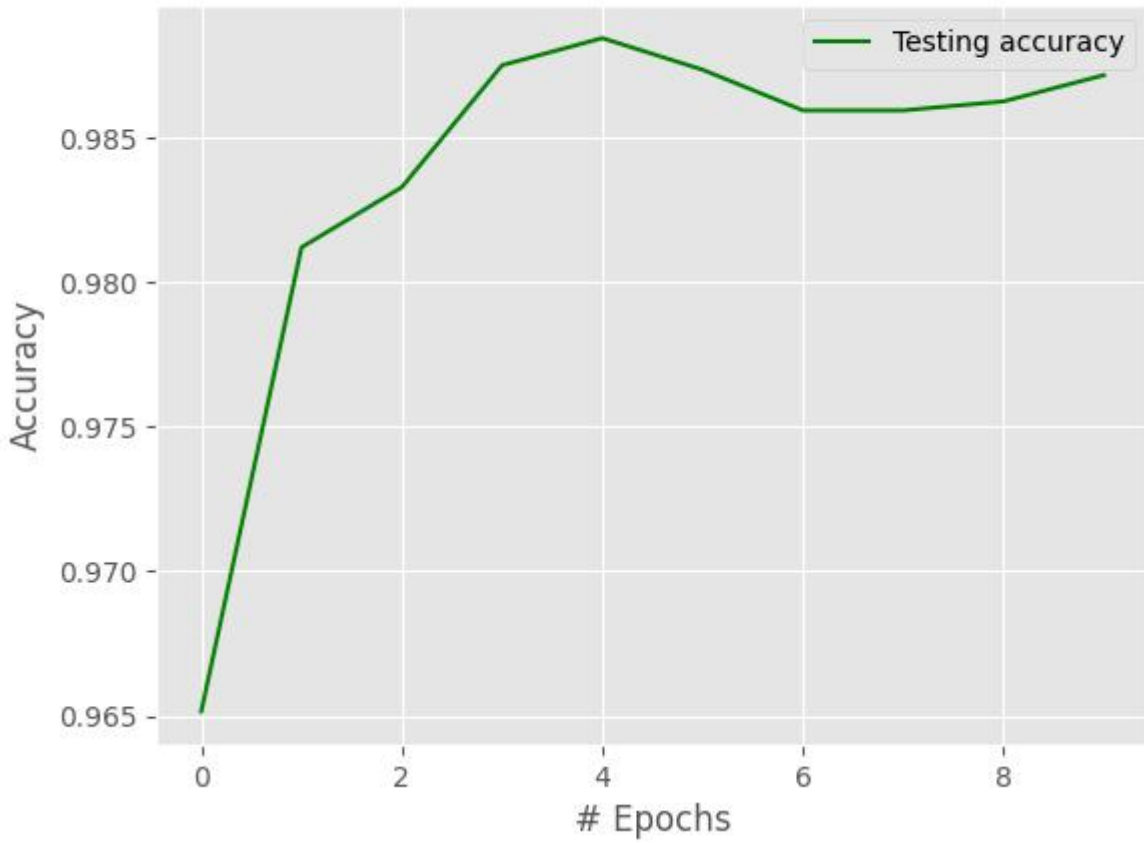
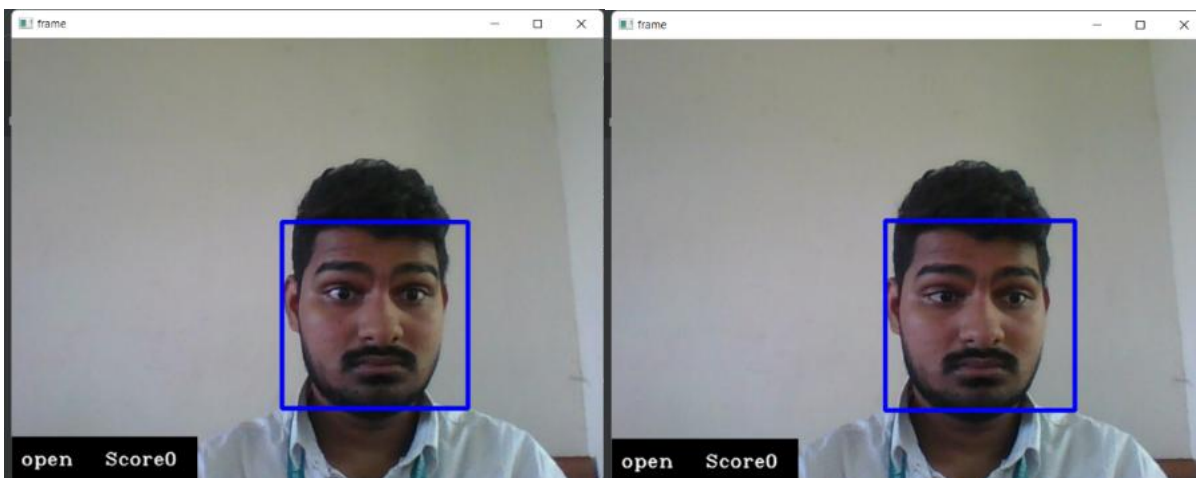


Figure 9: graph for accuracy



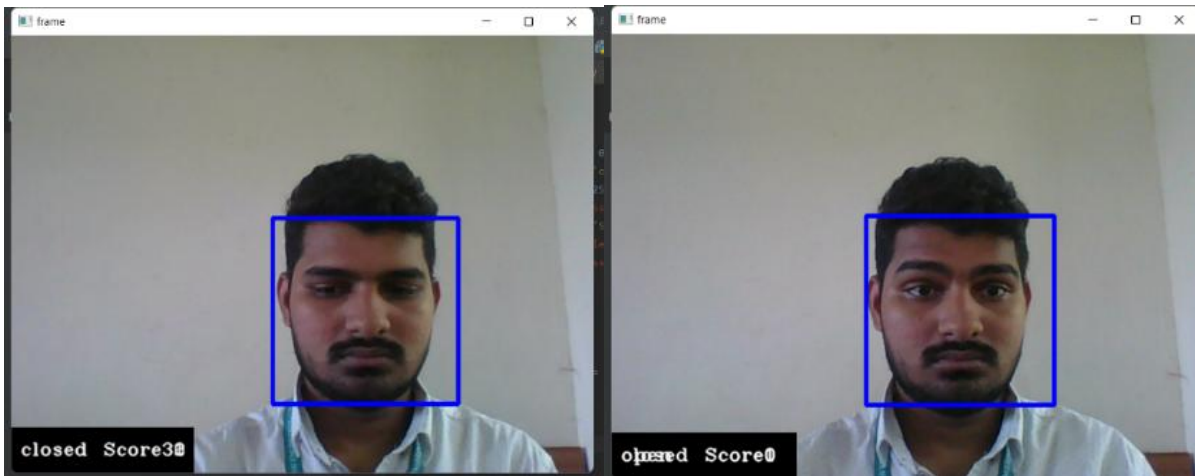


Figure 10: Output of all cases

Output of all the consecutive cases is given in figure 10.

9. CONCLUSION

We implemented a drowsiness detection application intended to reduce road accidents caused by driver fatigue. The proposed method uses a CNN-based transfer-learning algorithm to train the selected dataset and OpenCV for testing. It classifies a person as drowsy or alert. Future work could integrate this approach into a real-time device installed in a vehicle to monitor the driver's state.

REFERENCES

- Gwak, J., Hirao, A., Shino, M: An investigation of early detection of driver drowsiness using ensemble machine learning based on hybrid sensing. *Appl. Sci.* 10(8), 2890 (2020). <https://doi.org/10.3390/app10082890>
- You, F., Li, X., Gong, Y., Wang, H., Li, H.: A real-time driving drowsiness detection algorithm with individual differences consideration. *IEEE Access* 7, 179396–179408 (2019). <https://doi.org/10.1109/access.2019.2958667>
- Mehta, S., Dadhich, S., Gumber, S., Bhatt, A.J.: Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio. *SSRN Electron. J.* (2019). <https://doi.org/10.2139/ssrn.3356401>
- Song, F., Tan, X., Liu, X., Chen, S.: Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. *Pattern Recogn.* (2014).
- Savas, B.K., Becerikli, Y.: Real time driver fatigue detection system based on multi-task ConNN. *IEEE Access* 8, 12491–12498 (2020). <https://doi.org/10.1109/access.2020.2963960>

Bavkar, S., Iyer, B., Deosarkar, S.: Rapid screening of alcoholism: an EEG based optimal channel selection approach. IEEE Access 7, 99670-99682 (2019).

<https://doi.org/10.1109/ACCESS.2019.2927267>

Sathasivam, S., Mahamad, A.K., Saon, S., Sidek, A., Som, M.M., Ameen, H.A.: Drowsiness detection system using eye aspect ratio technique. In 2020 IEEE Student Conference on Research and Development (SCORED) (2020). <https://doi.org/10.1109/scored50371.2020.9251035>

Bavkar, S., Iyer, B., Deosarkar, S.: BPSO based method for screening of alcoholism. In: Kumar, A., Mozar, S. (eds.) ICCCE 2019. Lecture Notes in Electrical Engineering, vol. 570, pp. 47-53. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-8715-9_6

Bavkar, S., Iyer, B., Deosarkar, S.: Optimal EEG channels selection for alcoholism screening using EMD domain statistical features and harmony search algorithm. Biocybern. Biomed. Eng. 41(1), 83-96 (2021)

K. Vijay, R. Vijayakumar, P. Sivaranjani, and R. Logeshwari, "Scratch detection in cars using mask region convolution neural networks," Adv. Parallel Comput., vol. 37, pp. 575-581, 2020.

Deshpande, P., Iyer, B.: Research directions in the internet of every things (IoET). In: 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, pp. 1353-1357 (2017). <https://doi.org/10.1109/CCAA.2017.8230008>

Deshmukh, D., Iyer, B.: Design of IPSec virtual private network for remote access. In: 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, pp. 716-719 (2017). <https://doi.org/10.1109/CCAA.2017.8229894>

Carmel, M. B. M. J., S. Ravikumar, A. Muhammad, K. V. Dhilip, K. K. Antony, and G. Arulkumaran. 2022. "Linguistic Analysis of Hindi-English Mixed Tweets for Depression Detection." *Journal of Mathematics* 2022. <https://doi.org/10.1155/2022/3225920>.

Kepesiova, Z., Ciganek, J., Kozak, S.: Driver drowsiness detection using convolutional neural networks. In: 2020 Cybernetics & Informatics (K&I) (2020). <https://doi.org/10.1109/ki48306.2020.9039851>

Abtahi, S., Omidyeganeh, M., Shirmohammadi, S., Hariri, B.: YawDD: yawning detection dataset. IEEE Dataport (2020). <https://doi.org/10.21227/e1qm-hb90>.

Prithi S, Kishore A, Harinath D, Harini S "Incident Detection and Automatic Alert System", International Journal of Science, Engineering and Technology, Vol.8, Issue.2, 2020

Bhavani, M., Shrijeeth, S., Rohit, M., Sanjeev Krishnan, R., Sharveshwaran, R., "A detailed study on sentimental analysis using Twitter data with an Improved deep learning model", Proceedings of the 5th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2021.

Kannan, E., S. Ravikumar, A. Anitha, S. A. P. Kumar, and M. Vijayasathy. 2021. "Analyzing Uncertainty in Cardiotocogram Data for the Prediction of Fetal Risks Based on Machine Learning Techniques Using Rough Set." *Journal of Ambient Intelligence and Humanized Computing*.
<https://doi.org/10.1007/s12652-020-02803-4>.

Nilaiswariya, R., J. Manikandan, and P. Hemalatha. "Improving Scalability And Security Medical Dataset Using Recurrent Neural Network And Blockchain Technology." In 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1-6. IEEE, 2021.

R, Babu, Kannappan Jayashree, Krishna Brahmadesam Viswanathan, and Vijay K. 'An Efficient Spam Detector Model for Accurate Categorization of Spam Tweets Using Quantum Chaotic Optimization-Based Stacked Recurrent Network'. *Nonlinear Dynamics* 111, no. 19 (October 2023): 18523-40.
<https://doi.org/10.1007/s11071-023-08697-z>.