



PERFORMANCE ANALYSIS OF DRIVER DROWSINESS DETECTION USING SVM AND CNN

¹Dr. N. Radha, ²Ms. Mirunaalini K J

¹Associate Professor, ²Student

¹M.Sc Data Analytics,

¹PSGR Krishnammal College for Women, Coimbatore, India

Abstract - Road accidents have become a common phenomenon in these modern days. Reasons for road accidents are many. Driver drowsiness can be considered one of the major distractions that may lead to a road accident. To reduce the frequency of road accidents driver drowsiness system is made. This system is a safety alarm system that indicators the driver every time the driver feels drowsy. The eye movement of the driver is monitored and each time the driver feels asleep or closes their eye for a certain frame then it signals the driver with the help of an alarm. This is carried out by means of the usage of a machine learning and deep learning model to identify the face of the driver using a camera and analyze the state of the driving force. Two different algorithms are used to raise the detection accuracy, namely, Support Vector Machine (SVM), and Convolution Neural Network (CNN). SVM uses an Eye Aspect Ratio (EAR) for training the model. CNN uses the MLR eye dataset is used to train the model. The findings of this study show that CNN achieved the best result compared to SVM.

Index Terms – Deep Learning, Machine Learning, Convolution Neural Network, Support Vector Machine

I. INTRODUCTION

One of the most common causes of accidents is driver drowsiness and weariness. Every year, the number of people dying in accidents increases around the world. This article aims to reduce the frequency of accidents caused by driver weariness and drowsiness. This system is used to protect the driver by implementing a Drowsiness Detection System that detects whether the person's eyes are closed or open. Additionally, if a person's eyes are closed for specific frames, the system will notify them by ringing an alert sound.

II. BACKGROUND STUDY

When measuring the drowsiness level of a driver, there are two different approaches according to the origin of the data used for this measuring. On the one hand, there are systems that monitor the vehicle state to assess the fatigue of the driver, while on the other hand, there are systems that use parameters obtained from the own driver.

Systems focused on the vehicle

Among works that focus on the analysis of the vehicle state and its relation to fatigue, the most common measures that are studied are steering wheel behaviors or lane departures [2 - 4]. Other parameters of the car are used, such as the vehicle position or the steering wheel angle, and they perform data fusion on multiple measures to achieve a more reliable system. However, even if the diminishing performance over skill-based tasks by the driver can actually be a consequence of drowsiness, it appears at a later stage and cannot be used to detect the early symptoms of fatigue [5].

Systems focused on the driver

One of the most reliable ways of estimating fatigue is by using electroencephalograms (EEG) in combination with electrooculograms (EOG) [6], but in real driving environments, these kinds of systems are usually rejected by drivers. Their main drawback is that they require that the driver has attached electrodes around the eyes and over the head, which makes them intrusive systems that produce discomfort and rejection by drivers.

Because of this limitation, the most used fatigue detection systems are those in which the driver's state is detected through a camera placed on the vehicle that takes images of the driver. In this work, we will focus on the detection of the early symptoms of drowsiness by using the driver's state.

Behavioral measures:

Facial expressions-based drowsiness detection makes use of computer vision to detect and recognize facial motion and changes in appearance during drowsiness [7]. It accepts a stream of input images from a camera in front of the driver and passes these images through the four main image processing stages: face detection and tracking, feature extraction, feature selection, and classification.

III. METHODOLOGY

3.1 DATA MODELLING

This dataset is just one part of The MRL Eye Dataset, the large-scale dataset of human eye images. It's set up for classification work. This collection contains low and high-quality infrared photos acquired in various lighting circumstances and by various instruments. The dataset can be used to test a variety of features or trainable classifiers. The photos are separated into numerous categories to facilitate comparing algorithms easier, and they are also excellent for training and testing classifiers.

3.2 DATASET DESCRIPTION

This dataset contains two classes. One is eye open and the other one is eye closed. This dataset is used for training the model. And the real-time video is used to detect whether the driver is drowsy or not by using the trained model.

Subject ID: This dataset contains 37 individual eye images which comprise 33 men and 4 women.

Image ID: The dataset consists of 296 images

gender [0 - man, 1 - woman] The dataset contains the information about gender for each image (man, woman)

glasses [0 - no, 1 - yes]: The information if the eye image contains glasses is also provided for each image (with and without the glasses)

eye state [0 - closed, 1 - open]: This property contains the information about two eye states (open, close)

reflections [0 - none, 1 - small, 2 - big]: Three reflection states based on the size of reflections (none, small, and big reflections)

lighting conditions [0 - bad, 1 - good]: Each image has two states (bad, good) based on the amount of light during capturing the videos

sensor ID [01 - RealSense, 02 - IDS, 03 - Aptina]: At this moment, the dataset contains the images captured by three different sensors (Intel RealSense RS 300 sensor with 640 x 480 resolution, IDS Imaging sensor with 1280 x 1024 resolution, and Aptina sensor with 752 x 480 resolution) **Example:** s0028 00001 0 0 0 0 0 0 1



Figure 3.2.1
Sample eye closed image



Figure 3.2.2
Sample eye opened image

3.3 PROPOSED FRAMEWORK

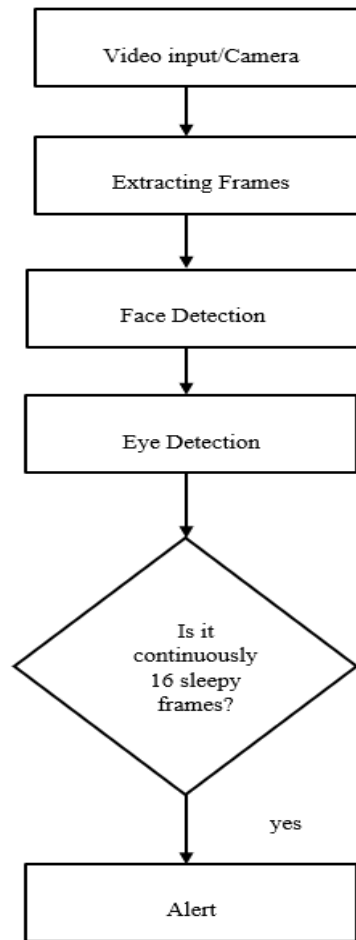


Figure 3.3.1 System Flow Diagram

3.4 PROCESS FLOW

This work first trains the model with an input image dataset using the following algorithms.

Proposed system algorithms

- i) The Viola-Jones face detection algorithm is used to detect the face of the images and given as input to the Viola-Jones eye detection algorithm.
- ii) Once the face is detected, the Viola-Jones eye detection algorithm is used to extract the eye region from the facial images and given as input to CNN.
- iii) CNN with four convolutional layers is used to extract the deep features and those features are passed to a fully connected layer.
- iv) Softmax layer in CNN classify the images into sleepy or non-sleepy images.

Face detection and eye region extraction

The whole face region may not be required to detect the drowsiness but only the eyes region is enough for detecting drowsiness. In the first step by using the Viola-Jones face detection algorithm, the face is detected from the images. Once the face is detected, the Viola-Jones eye detection algorithm is used to extract the eye region from the facial images. For face detection, the Viola-Jones algorithm has three techniques those are Haar-like features, Ada boost, and Cascade classifier. In this work, the Viola-Jones object detection algorithm with Haar cascade classifier was used and implemented using OPEN CV with python. Haar cascade classifier uses Haar features for detecting the face from images.

Feature extraction and classification

Feature extraction is one type of dimensionality reduction where useful parts of an image are represented as a feature vector. In this project features from the eye, and region images are extracted using a Convolutional Neural Network (CNN).

Convolutional neural network

A convolutional neural network (CNN) is used in the proposed system for the detection of driver drowsiness. Since a feature vector is needed for each drowsy image to compare with existing features in a database to detect either drowsy or not. Usually, CNNs require fixed-size images as input so pre-processing is required. The pre-processing includes extracting the keyframes from video based on temporal changes and store in the database. From these stored images, feature vectors are generated in convolution layers of CNN. These feature vectors are then used for detecting the driver's drowsiness. CNN has layers like convolutional layers, pooling (max, min, and average) layers, ReLU layers, and fully-connected layers. The convolution layer is having kernels (filters) And each kernel has width, depth, and height. This layer produces the feature maps as a result of calculating the scalar product between the kernels and local regions of the image. CNN uses pooling layers (Max or Average) to minimize the size of the feature maps to speed up calculations. In this layer, the input image is divided into different regions then operations are performed on each region. In Max Pooling, a maximum value is selected for each region and placed in the corresponding place in the output. ReLU

(Rectified Linear Units) is a nonlinear layer. The ReLU layer applies the max function on all the values in the input data and changes all the negative values to zero. The following equation shows the ReLU activation function.

$$F(x)=\max(0,x) \quad (1)$$

The fully-connected layers are used to produce class scores from the activations which are used for classification.

In the proposed method, 2 convolutional layers and one fully connected layer are used. Extracted key images with the size of 24 X 24 are passed as input to the convolution layer-1 (Conv2d_1). In Conv2d_1 input image is convolved with 32 filters of size 3x3. After convolution, batch Normalization, non-linear transformation ReLU, and Max pooling over 1 × 1 cells are included in the architecture, which is followed by dropout with 0.25%. The output of convolution layer-1 is fed into the convolution layer-2(Conv2d_2). In Conv2d_2, input is convolved with 32 filters with size 3x3 each. After convolution, batch Normalization, non-linear transformation ReLU, MaxPooling over 2 × 2 cells with stride 2 followed by dropout with 0.25% applied. Conv2d_2 required 9248 parameters. The output of convolution layer-2 is fed into a fully connected layer. The proposed CNN model required trainable parameters. The output of the classifier is two-state, so the output layer has only two outputs. Adam's method is used for Optimization. Here softmax classifier is used for classification. In this proposed CNN framework, the outputs of a fully connected layer are the deep features retrieved from input eye images. The final outputs can be the linear combinations of the deep features.

Support Vector Machine (SVM)

In this algorithm, our modeling is based on a metric called Eye Aspect Ratio. The EAR depends on the computation of the distance ratio between previously specified facial landmarks of the driver's eyes. It reflects the eye's openness degree allowing us to detect the initial signs of drowsiness. A very low EAR value means that the eye is closed. Eyeblink is a quick closing and reopening of the eye. Each person has a slightly different blinking pattern. It mainly differs in the closing and opening speed, blink duration, and the degree to which the eye is closed. Normal eye blink lasts approximately 100 to 400 msec. In order to detect the different blinks for each person and since one EAR value per frame will not recognize the eye blinks correctly, the proposal is to train a classifier that uses a large temporal window of frames as an input. For every 30 f/s videos, we computed the EAR of the Nth frame, along with the EAR for N-7 and N+7 frames. Then, by concatenating these EARs, a 15-dimensional feature vector is formed for each frame. This vector is then used as an input to the SVM.

The linear SVM classifier is a common classifier used in many other drowsiness detection studies; thus, it was used to test our data. Finally, the CNN classifier was used to apply deep learning to test if the model will produce higher results. Overall, all of the classifiers were used to classify the driver's eyes as open or closed. Generally, eye closures that exceed 500 msec indicate entering a drowsy state. However, since blink duration differs from one person to another, It decided to set the drowsiness threshold to 1 second, which is equivalent to getting fifteen consecutive predictions of closed eyes. Thus, if fifteen consecutive predictions were labelled "closed eye," an alarm will raise, indicating that the driver is drowsy.

Table 3.4.1 Performance Comparison Table

Algorithm	Accuracy
Support Vector Machine	93.6 %
Convolution Neural Network	94.03

IV. MODEL EVALUATION

4.1 EXPERIMENTS

In this paper, there are two types of experiments were conducted. The first experiment was performed on the image dataset. In the second type, the experiment was performed on video. To conduct the first type of experiment, download the dataset from Kaggle. In these, 120 drowsy images and 120 non-drowsy images are used for training. And 120 images are used for testing out of which 60 images are drowsy images and another 60 images are non-drowsy images the proposed model has achieved an accuracy of 94.03% on the test dataset. The accuracy of the proposed model after 20 epochs with batch size 3. In the second type of experiment, train the model with 240 samples. During the testing phase, capture the video frames through the camera and alert with an alarm when the model predicts a drowsy output state continuously. Static images are used for training but during the testing phase, keyframes are extracted from continuous video and tested against the trained static images. The experiments shown in Figure 4.1 indicate non-drowsy and Figure 4.2 indicate drowsy.



Figure 4.1 Eye Open



Figure 4.2 Eye Close

V. CONCLUSION

In this paper, a model is proposed for driver drowsiness detection based on eye state. This determines the state of the eye that is drowsy or non-drowsy and alert with an alarm when the state of the eye is drowsy. Face and eye regions are detected using the Viola-Jones detection algorithm. A deep convolution neural network is developed to extract features and used for the learning phase. A SoftMax layer in the Convolution Neural Network (CNN) classifier is used to classify the driver as sleep or non-sleep. The proposed system achieved 96.03% accuracy. The proposed system effectively identifies the state of the driver and alerts with an alarm when the model predicts a drowsy output state continuously.

The proposed driver drowsiness detection system using SVM has some limitations. Since only the EAR is used as a drowsiness indicator, false alarms might rise when the driver is laughing or yawning. The proposed system achieved 93.6% accuracy. From the experiments, it was observed that Convolution Neural Network (CNN) performs well for detecting driver drowsiness detection than the Support Vector Machine (SVM).

REFERENCES

1. Amodio A, Ermidoro M, Maggi D, Formentin S, Savaresi S.M. (2018). Automatic detection of driver impairment based on pupillary light reflex. *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-11.
2. Krajewski J, Sommer D, Trutschel U, Edwards D, Golz M. Steering Wheel Behavior-Based Estimation of Fatigue. In *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Big Sky Resort, Big Sky, MT, USA, 22–25, 2009; pp. 118–124.
3. Friedrichs F, Yang B, Drowsiness monitoring by steering and lane data-based features under real driving conditions. In *Proceedings of the 2010 18th European Signal Processing Conference*, Aalborg, Denmark, 23–27, 2010; pp. 209–213.
4. McDonald A.D, Schwarz C, Lee J.D, Brown T.L. Real-Time Detection of Drowsiness Related Lane Departures Using Steering Wheel Angle. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2012*; 56, 2201-2205.
5. Yang J.H, Mao Z.H, Tijerina L, Pilutti T, Coughlin J.F, Feron E, Detection of Driver Fatigue Caused by Sleep Deprivation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*2009; 39, 694–705.
6. Sommer D, Golz M, Evaluation of PERCLOS based current fatigue monitoring technologies. In *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires, Argentina, 2010; pp. 4456–4459.
7. S. Ghosh, T. Nandy, and N. Manna, Real-Time Eye Detection and Tracking Method for Driver Assistance System, *Advance in Electrical Engineering*, 2015.
8. Viola P, Jones M.J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2); pp. 137-154.
9. Liang Y.L, Reyes M.L, Lee J.D. Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation* 2007, 8(2); pp. 340–350.
10. Hu S Zheng G. Driver drowsiness detection with eyelid-related parameters by support vector machine. *Expert Syst Appl* 2009; 36(4); pp 7651–7658.
11. O'Shea K, Nash R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.
12. Lee K, Yoon H, Song J, Park K (2018). Convolutional neural network-based classification of driver's emotion during aggressive and smooth driving using multi-modal camera sensors. *Sensors*, 18(4): 957.