e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:06/June-2022 Impact Factor- 6.752 www.irjmets.com

DRIVER DROWSINESS DETECTION USING HAAR CASCADES

Shubham Chowdhury[*1], Dr. Jibrael Jos[*2]

[*1,2]Department Of Data Science, Christ (Deemed To Be) University,

Lavasa, Pune, Maharashtra, India.

ABSTRACT

This document outlines the architecture and working of a driver drowsiness detection system built using HAAR cascades and convolutional neural networks, designed to provide an early- warning system for drowsy driving situations, thus preventing the occurrence of an accident. The implementation has been suggested as an additional safety mechanism for automobiles and provide a deterrent for road accidents caused by manual error. The facial feature detection is suggested using HAAR cascades, and the classification based on these features is suggested using convolutional neural networks.

Keywords: Driver Drowsiness, Convolutional Neural Networks, Safety Mechanism, HAAR Cascades.

I. INTRODUCTION

According to the National Highway Traffic Safety Administration, every year about 1,00,000 police-reported crashes are found to be caused due to drowsy driving. These crashes result in more than 1550 fatalities and 71,000 injuries. The real number may be much higher, however, as it is difficult to determine whether a driver was drowsy at the time of a crash, and if this condition primarily caused the crash.

In order to make the driver aware of such situations before they occur, a system is required to make sure that the driver is alerted when symptoms of getting drowsy are detected and thus, a mishap can be avoided. Several approaches are proposed to help with this detection. Some of them involve tracking biological markers of the driver to predict dangerous behavior before it occurs, whereas some methods focus on detecting the unsafe behavior as soon as it can be observed physically and avoid it cascading to a crash.

One of the methods is detecting biological factors that are triggered when a person is sleepy: EEG/ ECG/ EOG data can be collected and correlated with driver awareness states. However, this approach is intrusive, as several sensors need to be connected to the driver to detect these signals and neuro-indicators. This approach, therefore, is not highly suitable for everyday driver applications.

Another approach that is suggested is detection of lane changes with the help of external cameras on the vehicle. When a driver is drowsy, he/ she has difficulty maintaining a straight line along a lane. Deviations from the extreme lanes outwards, and unindicated changes in lane, can therefore, be possible triggers for this system. This system might be useful in cases where these cameras are also put to some alternate uses alongside this application, like self-driving or navigation, and in situations where drivers are expected to follow stringent lane discipline, e.g., on speedways and highways.

The approach that is considered for this experiment is changes in facial expressions, by observing the driver inside the vehicle cabin with the use of cameras. We aim to predict the drowsy state of a driver on the road by identifying facial indicators of their sleep- deprived or tired state. This approach is considered for the purpose of this project for the following reasons:

- A camera is non- intrusive, compared to sensors and headsets needed to be put on when driving,
- The system is fairly easy to install in new and old vehicles alike,
- This system can be readjusted to be suitable for different environments where detection and accuracy may be affected (such as use of infrared cameras in predominantly dark cabins),
- The model can be retrained for varying facial features, and can be improved upon with suitable training data,
- It does not require very high computational capabilities to be deployed, so it can be integrated with the electronic control unit or the software processor of the vehicles.

The primary facial indicators considered for this approach are closed eyes, and yawning. These are called bio-indicators, which are manifestations that can be observed and measured physically, as an indication of the subject's fatigue. In addition to these, there are several bio- indicators, like changes in heart rate, heart rate

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:06/June-2022 Impact Factor- 6.752 www.irjmets.com

variability, and respiratory rate, which can be used to measure fatigue. These features, however, require wearable sensors to be worn by the subject for measurement. Since the target is to analyze the drowsiness with as little intrusion as possible, use of the first two indicators was the optimal way to go.

The idea of real- world implementation of this system is as follows: a camera is set up within the cabin of the vehicle, such that the driver's face is clearly visible in the image, while keeping the rest of the vehicle's occupants out of frame. A real- time feed of this camera can be used to monitor the facial features of the driver using the drowsiness detection system. In case the driver is found to be drowsy, an alert is sent to the driver to caution them of their hazardous state. This can be done by one of a combination of these approaches: a voice alert suggesting them to take a break from driving; haptic feedback to the steering wheel; haptic feedback to the driver's seat.

For this purpose, the features first need to be detected as per relevance, like eyes, mouth, eyebrows etc. Based on these features, the model can be trained to detect drowsiness from the provided dataset, which can be images/ videos of drivers in different states of alertness, with the features being the inputs and the drowsiness state as the output.

II. METHODOLOGY

This chapter discusses the functional requirements which include, identifying the best parameters, and correct weights, learning rate etc. and non-functional requirements such as scalability and reliability.

The proposed algorithm consists of the following steps:

1. Obtaining a feed of the driver's face, which in this case is done by capturing real- time images of the driver taken with a camera attached within the cabin of the car facing the driver,
2. Detection of the driver's face from the obtained images,
3. Predicting the landmarks of important regions in the detected face, and
4. Comparing the ratios of these detected features with pre- trained models to classify them as symptoms of drowsiness,
5. Providing an alert on exceeding a particular threshold

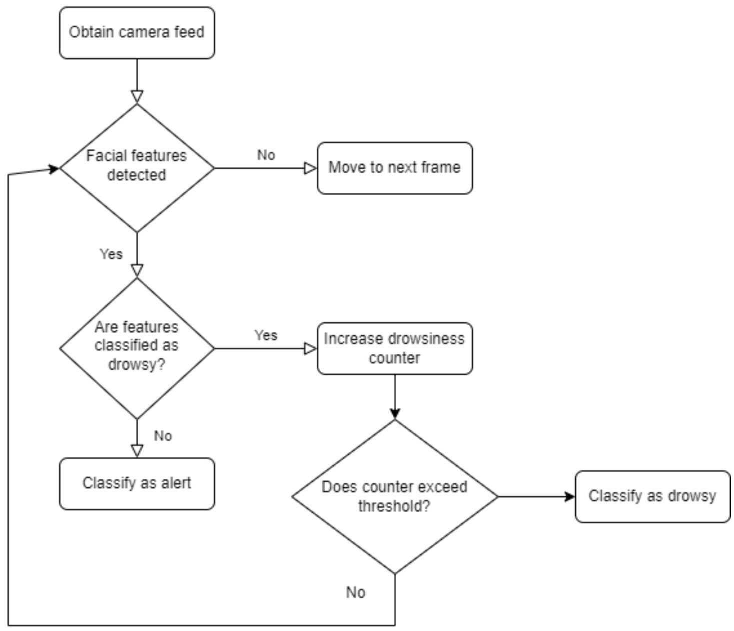

Figure 1: Process flow block diagram

The facial features are detected by the implementation of HAAR cascade files, which are xml files for HAAR cascade operations. The convolution of these cascades over the entire image results in the detection of these features across the entire image, the combination of which make up the features we are looking to identify. The feature set is reduced from the list of all features by using a boosting technique called Adaboost, in which a combination of these features is used to create weak learners, designed in a way that they would misclassify a minimum number of images.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:06/June-2022 Impact Factor- 6.752 www.irjmets.com

These features are applied to sections of the image that are 24 pixels by 24 pixels in size, in stages of increasing complexity. The stages with less complex features remove most of the images with no facial features, thus reducing the false negative ratio and providing a better dataset for the detection of complex features. This significantly reduces the number of calculations required for detecting these more complex features from within the image.

Network Architecture

The architecture for the CNN consists of these layers:

- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 64 nodes, kernel size 3
- Fully connected layer; 128 nodes

The final layer is connected to the output layer by a fully connected layer with 2 nodes. In all the convolutional layers, a Relu activation function is used and the output layer uses a Softmax activation function. Once the features are detected in the face that indicate drowsiness, a threshold is decided for classifying the driver as drowsy. Whenever the threshold is exceeded, in this case, when the eyes are closed for over 15 consecutive frames, the driver is classified as drowsy.

III. MODELING AND ANALYSIS

The "ULg Multimodality Drowsiness Database", also called DROZY, is a database containing various types of drowsiness-related data (signals, images, etc.) and intended to help researchers to carry out experiments, and to develop and evaluate systems (i.e., algorithms), in the area of drowsiness monitoring. The data in the DROZY database was collected from 14 healthy and young subjects (consisting of 3 males, and 11 females) who consented to be subjected to 3 successive 10-min psychomotor vigilance tests (PVTs) in conditions of increasing sleep deprivation, induced by acute, prolonged waking.

ImageNet is an image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images. The project has been instrumental in advancing computer vision and deep learning research. The data is available for free to researchers for non-commercial use. Imagenet has been used to obtain the images required for training the classification model. Over 7000 images were extracted from the dataset.

Custom HAAR cascades were created for detecting the facial features. 250 positive images and 400 negative images were used for generating the xml file, achieving an accuracy of 78% on testing with a new dataset of face portrait images. Some misclassifications occurred when multiple faces were detected in the image, some faces were partially visible, or certain patterns and textures were present as part of the background. Accuracy was slightly increased by adjusting the scale factor to lie between 1.5 and 3. Values higher than that led to faster detection, with the drawback of missing some faces from detection completely.

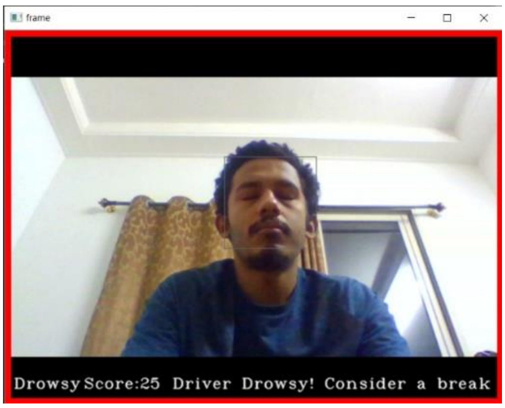

Figure 2: Sample output – drowsy

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:06/June-2022 Impact Factor- 6.752 www.irjmets.com

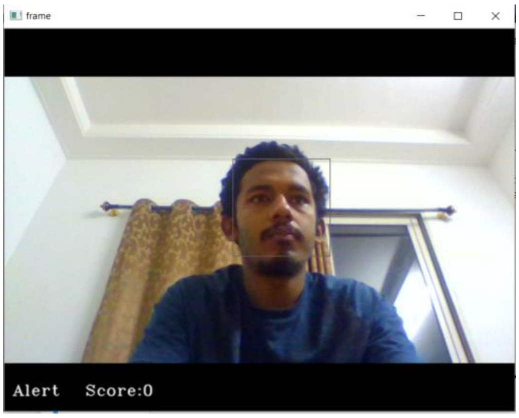

Figure 3: Sample output – alert

Another parameter whose change had significant impact on the detection performance is the value of minNeighbors parameter, which is essentially the number of neighbors that each candidate rectangle should have to retain it. Higher values for this parameter were found to lead to lesser detections, but with higher accuracy.

IV. RESULTS AND DISCUSSION

The highest levels of accuracy for HAAR cascade feature detection were achieved with minimum misclassifications when scaling factor was set to 2, and minNeighbors to 6. It is to be noted, however, that these were the optimal value for the resolution of images used for testing this model, and may vary for different setups and testing images.

After the feature detection was finalized, the next part was to figure out the best models for classification of the driver as drowsy or alert, based on these features.

Various modelling techniques were employed after the features were extracted and normalized, including logistic regression, Naïve Bayes, Decision Tree, Random Forest, CNN, and k-nearest neighbor. Although CNN was the most preferrable choice for feature detection, k-nearest neighbor is found to perform most accurately for generating the classification model with an accuracy of 71.6%.

K- nearest neighbor (k = 30)	71.6
Decision Tree	68.9
Random Forest	67.6
XGBoost	65.2
CNN	61.1
Naïve Bayes	52.7

Figure 4: Accuracy of different models

However, in this case, the recall values were found to be higher than desired, as this would lead to false negatives, i.e., drowsy drivers being classified as alert.

Therefore, CNN was used as the classifier, as even with lower accuracy, precision and recall values were more desirable than with k-nearest neighbor.

V. CONCLUSION

The automobile industry is a rapidly evolving industry, with new technologies and drivetrains being developed, tested, and manufactured every day. One of the primary regions that buyers focus on, apart from performance, is safety. The methods discussed in this paper are a step in making automobiles a safer mode of transport in the future. The technologies suggested can make their way to future automobiles and their implementation can have a considerable effect on road safety.

The classification model was built for the purposes suited for this implementation. Given the conditions, the accuracy achieved by using CNN was relatively high with less false negatives reported in a new dataset. The

IRJMETS

e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:04/Issue:06/June-2022 Impact Factor- 6.752 www.irjmets.com

feature detection methods used were found to be accurate under varying conditions of lighting and head tilt, except in extremely low light conditions. An alternate approach of using infrared illumination in such cases has been discussed. Using k- nearest neighbor was ruled out, with high rate of false negative detection being a critical drawback of the model. Since drowsy drivers being detected as alert was a violation of the purpose of this project, the parameters of using this model were unsuitable for this application.

VI. REFERENCES

[1] Paul Viola & Michael Jones, Rapid object detection using a boosted cascade of simple features. MITSUBISHI ELECTRIC RESEARCH LABORATORIES, 2001.

[2] P. Viola and M. M. Jones, "Robust Real-Time Face Detection," International Journal Of Computer Vision, 2004.

[3] Zuopeng Zhao, Nana Zhou, Lan Zhang, Hualin Yan, Yi Xu, and Zhongxin Zhang, "Driver Fatigue Detection Based on Convolutional Neural Networks Using EM-CNN", 2020.

[4] R. Onken, "DAISY, an adaptive, knowledge-based driver monitoring and warning system," in Proceedings of the Intelligent Vehicles Symposium, 1994.

[5] S. Singh and N. Papanikolopoulos, "Monitoring driver fatigue using facial analysis techniques," in Intelligent Transportation Systems, 1999.

[6] E. M. Ayoob, A. Steinfeld and R. Grace, "Identification of an "Appropriate" Drowsy Driver Detection Interface for Commercial Vehicle Operations," in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2003.

[7] L. Fletcher, L. Petersson and A. Zelinsky, "Driver assistance systems based on vision in and out of vehicles," in Proceedings of the IEEE Intelligent Vehicles Symposium, 2003.

[8] Q. Ji, Z. Zhu and P. Lan, "Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue," IEEE Transactions on Vehicular Technology, 2004.

[9] P. Smith, N. V. Lobo and M. Shah, "Determining Driver Visual Attention With One Camera," IEEE Transactions On Intelligent Transportation Systems, December 2003.

[10] Ralph Oyini Mbouna, Seong G. Kong and Myung-Geun Chun, "Visual Analysis of Eye State and Head Pose for Driver Alertness Monitoring", IEEE, 2013.

[11] Z. Zhu and Q. Ji, "Real Time and Non-intrusive Driver Fatigue Monitoring," in IEEE Conference on Intelligent Transportation Systems, Washington D.C., USA, 2004.

[12] R. Hanowski, J. Hickman, M. Blanco and G. Fitch, "Long-haul truck driving and traffic safety: Studying drowsiness and truck driving safety using a naturalistic driving method.," Verster, J.C. (Ed.) Sleep, Sleepiness and Traffic Safety. Hauppauge, NY: Nova Science Publishers., 2010.

[13] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea and M. E. Lopez, "Real-Time System for Monitoring Driver Vigilance," IEEE Transactions on Intelligent Transportation Systems, March 2006.

[14] M. Eriksson and N. P. Papanikolopoulos, "Driver fatigue: a vision-based approach to automatic diagnosis," Transportation Research Part C: Emerging Technologies, 2001.

[15] M. Sacco and R. Farrugia, "Driver fatigue monitoring system using Support Vector Machines," in 5th International Symposium on Communications Control and Signal Processing, 2012.

[16] G.C.Feng and P.C.Yuen, "Multi-cues eye detection using gray intensity image," Pattern Recognition, Elsevier, January 2000.

[17] Y. Freund and R.E.Schaphire, "A decision-theoretic generalization of on-line learning and an application to boosting," Computational Learning Theory: Eurocolt 95, Springer-Verlag, 1995.

[18] Anirban Dasgupta, Anjith George, S L Happy and Aurobinda Routray, "A Vision Based System for Monitoring the Loss of Attention in Automotive Drivers", IEEE, 2013.

[19] Karamjeet Singh and Rupinder Kaur, "Physical and Physiological Drowsiness Detection Methods", International Journal of IT, Engineering and Applied Sciences Research (IJIEASR), 2013.