Driver’s Drowsiness Detection Based on Facial Multi-Feature Fusion

V. Venkata Sai Vardhan1, N. Ritish Kumar Reddy 2, K. Jaya Surya3, J. Uday Kiran4, Ashwani Kumar5,

1,5Department of Computer Science and Engineering, Sreyas Institute of Engineering and Technology, Hyderabad, India., ashwani.kumarcse@gmail.com

Abstract. Among the leading causes of traffic accidents and deaths is drowsy driving. As a result, detecting and indicating driver drowsiness is an important research subject. The majority of existing approaches are based on automobiles, behavioral-based, or physiologically based. A few approaches are invasive and so the driver been interrupted, while others necessitate the use of pricey sensors and data processing. As a result, real time driver sleepiness detection system with adequate consistency and minimized cost has been established in this work. A webcam captures the video in the proposed system, and image processing techniques are used. Histogram of oriented gradients (HOG) which is available in Dlib toolkit, is used to recognize the driver's face in every frame. The landmarks on the face are identified by State Vector Machine, and the eye aspect ratio, mouth aspect ratio, are computed, and drowsiness is recognized using created adaptive thresholding based on their values. Offline implementations of machine learning techniques were also been made. In terms of accuracy and speed, we show that our algorithm outperforms existing fatigue methods. In Support, a sensitivity of 95.28 percent and a response rate of 100 percent were attained in Support Vector Machine.

1. Introduction

In recent decades there is a significant advancement in automobile industries and there are also road accidents that are rapidly increasing. Some of these accidents also happen due to the medical condition of the driver (i.e. Heart Attack etc.,) and majority of them are because of drowsiness. These accidents not only result in economic loss but may also result in physical injuries, which could result in permanent disability or even death. This condition is affects or more probable to the drivers who has sleep deprivation or sleeping disorders problems, drink alcohol’s, who drove for lengthy hours and long distances. Passengers in every country face the nightmare of drowsy drivers because drowsiness leads to lack of attention and concentration over steering of the vehicle and also drifting from the lanes. Every year due to fatigue there are substantial number of traffic accidents and results in an unusually high number of injuries and deaths each year, so it is a cloudy nightmare for passengers and other drivers too.

As a result, detecting and indicating driver weariness is a full-time job and due to its wide practical application, detecting and indicating driver fatigue is a hot topic of research. The emergency-response system, processing and acquisition systems are the three components of basic sleepiness detection system. The acquisition system captures a images from the video which consists face of human, which is then flows to the next block i.e.. processing block, where the images being processed online to predict the drowsiness.

The driver receives a warning or alert from the warning system if sleepiness is detected. In general, there are three sorts of approaches for detecting sleepy drivers: vehicle-based, behavioral-based, and physiological-based. A multitude of parameters like movements of wheel, acceleration and brake sequence, deviations in speed of vehicle, and variations in actual driving path and lane positions are used in a vehicle-based technique. In vehicle based methods mostly two commonly used methods they are...
Steering Wheel Movement and Standard deviation lane position. The SWM (steering wheel movement) can only work correctly in specific conditions and are heavily focused on the road's geometrical parameters and, to a smaller degree, on vehicle's kinetic properties. SDLP has the drawback of being entirely reliant on external elements such as traffic signs, climate, and brightness. Several researchers have discovered that vehicle-based measurements are a bad measure of drowsiness and therefore it has poor efficiency. Driver drowsiness is defined as the identification of any unusual changes in these parameters. Because the sensors are not linked to the driver, this is a non-intrusive measurement. To identify sleepiness, the behavioral-based technique analyses the driver's optical activity, such as eye blinking, eye shutting, yawning, head bending, and so on. These are the attributes to evaluate drowsiness. This is also a non-intrusive measurement since these characteristics are recognized using a basic webcam. Physiological signals such as Electrocardiogram (ECG), Electrooculogram (EOG), Electroencephalogram (EEG), heartbeat, blood pressure, etc being examined and monitored in the physiological-based technique, and the amount of sleepiness or exhaustion is determined based on these metrics. Because the sensors are linked to the driver, this is an invasive measurement that might distract the driver. EEG is by far the most frequently utilized physiological measurement. Whereas the EEG method appears like a viable method in confirming driver sleepiness, it's sometimes hard to set up in practice due to distortions created by body, head, and facial expressions, that could distort the Brain signals. The cost, complexity of building the systems, and size of the system will grow depending on the sensors utilized. But, adding more boundaries/features will enhance the system's accuracy to some amount. These criteria encourage us to create the least cost-based and real-time-based sleepiness detection system for drivers that is very accurate. As a result, we presented a system that uses a webcam that is capable of identifying drowsiness from a facial image employing simple image processing and ML strategies to enhance the system efficiency and transportable.

2. Related Work
In the paper written by Jayasenan J. S and Mrs. Smitha P. S says some approaches are broadly categorized into three types: 1) vehicle-based, 2) behavioral-based, and 3) physiological-based. In the first method, the drowsiness can be calculated using few measures like movements of a wheel, acceleration and brake sequence, deviations in the speed of the vehicle. In behavioral-based methods based upon the eye blinking, eye closing, yawning rates drowsiness is evaluated. In Physiological based methods, consider pulse, heartbeat rates are monitored to detect drowsiness. Kumar et al. presented object detection technique using SSD algorithm [6-9]. They have used machine learning based methods along with single SSMD detector algorithm to develop the model. In the paper which is written by Amir Jalilifard and Ednaldo Brigante Pizzolato. The aim of the paper is drowsiness detection using an efficient known algorithm, authors first used a short-time Fourier transform for distributing the power in the time-frequency space, the main reason to do so is to calculate the mean for the sub-bands. Out of all only 52 features are considered and an rf algorithm is imposed on these features to select only the best or more suited ones. this new feature is used to detect fatigue. Authors also claim that the system shows 91% accuracy.

3. Proposed Method
We propose a Driver’s Drowsiness Detection Algorithm in which the eyes and mouth positions are taken as features to detect the driver’s drowsiness. The face is detected in the frames using HOG and SVM to classify the components from the video which is in a live recording. The primary goal of this research paper is to develop a facial image descriptor and classification algorithm capable of detecting the drowsiness of drivers in different ambient conditions. PERCLOS is a fundamental measure used in evaluating the drowsiness of the driver, PERCLOS is a measure that counts the no. of times the driver closes and opens his/her eyes. But this can't be taken as the basis for concluding that the driver is drowsy, as there are other factors that the driver may close and open his/her eyes which depends upon the dust, intensity of lights. As our application is based on the behavioral characteristics of the driver we don't require any hardware that is attached to the driver. the behavioral characteristics like eye and mouth are observed using a camera and the rest of the process is done in the background which in no way distracts the driver. In this research paper, both accuracy and fastness are required at optimum levels as one without the other is a huge waste of time. the algorithm used in the proposed system is showing optimum results in both accuracy and speed.
The above block diagram shows how the proposed system operates. The whole system can be summarized as below using the above block diagram. The camera starts recording the video and it is given as an input to our first algorithm HOG, which is used in detecting the face, and then to our second algorithm SVM, which is used to detect the features of the face, the result after this two algorithm one after the other is used to calculate and compare with the threshold values based on which the status of the alarm is decided whether to on or off. Based on the above block diagram, our research work is divided into the following modules. Proposed system has observed to show approximately 95.28% accuracy.

1. Face Detection
2. Feature Detection
3. Drowsy State Identification
4. Alert System

3.1 Face Detection:
The working of our application starts with the recording of the driver. Detecting the face of the subject is our priority as this greatly affects the later process. Face detection is considered to be the first module that is implemented using the HOG algorithm. HOG is the short form for Histogram of Oriented Gradients, it is a feature descriptor that is used to identify the object as it is explained in the above chapter. So, HOG is used as an image descriptor in our application.

3.2 Feature Detection:
As we already mentioned that our application is based on behavioral characteristics, so features status are used to state whether the driver is drowsy or not. Feature detection is the module where we use our second algorithm, SVM. SVM stands for Support Vector Machine which is a classification algorithm used to detect features on the subjects face accurately using facial landmarks. This module is implemented on every frame which is outputted from the previous module.

3.3 Drowsy state identification:
This is the module where the calculations and comparisons are done on the results of the previous module. As we said the eyes and mouth status are taken into consideration in concluding whether the subject is fatigue. Two operations are carried out in this module which are described below,

3.4 Calculation:
In detecting a drowsy state we calculate the aspect ratio of the eyes and mouth which are explained in chapter 3. Since the aspect ratio is calculated separately for each eye, we take an average of the values of both eyes.

3.5 Comparison:
The calculated aspect ratio values are compared with the threshold to decide the status of the driver. The threshold values are given to the system or we can say they are taken by default. The calculated aspect ratio of the eye has to be greater than the threshold if not the application considers that the driver is drowsy. It is reverse in the case of mouth, the calculated aspect ratio of the mouth should be less than the threshold if not the application considers that the driver is yawning.

3.6 Alert system:
This is the module of various options. Alerting the driver is our intention in this module. Alerting the subject can be achieved in many ways from sending a message to taking control of the car. Our system
alters the driver by an alarm sound.

**Algorithm:**
The description of our proposed model is mentioned below:

**Input:** shape_predictor_68_face_landmarks  
**Output:** Estimating drowsiness of the driver and alerting them.

**Step 1:** The video is recorded from the camera that captures the whole face of the driver  
**Step 2:** From the recorded video we extract 2D frames.  
**Step 3:** The Frames or images extracted should be converted to grayscale image because it took less memory and processing speed is quick  
**Step 4:** The images generated are stored into an array using numpy module in python.  
**Step 5:** HOG is used to detect the faces in the extracted frames This will results in the only face in the given image  
**Step 6:** The output of HOG will be given to the trained SVM model to generate 68 facial landmarks coordinates on the face image and then it classifies the eyes and mouth,  
**Step 7:** We calculate the Aspect ratio of eyes and mouth.  
**Step 8:** Then Compare the calculated values with the Threshold value.  
**Step 9:** Based on the comparison then we conclude whether the person is drowsy or not.  
**Step 10:** Trigger the alert sound if fatigue is detected.

4. Methodology
Implementation is one the most important steps in any project life cycle where all of the effort in planning and gathering of information are concluded here. Though implementation is a crucial area it mostly depends on the planning and requirements of the project. Implementation can be simply defined as putting the plan into action or creating something that meets our requirements. Meeting requirements is where the heart of the implementation lies but that does mean that implementation is only coding, there is a lot more that has to be done to put the code or the application into use.

4.1 HOG:
Histogram of Oriented Gradients is one the ML algorithm which is used to identify objects in the input image. It is a feature descriptor algorithm used to identify the face in an image. To implement HOG we use the python Dlib module. The output of the HOG can be used well as they are incorrect format to apply classification or regression algorithm. Before going into how the hog works, we need to understand few terms  

4.2 Gradient:
Sudden change in the color like an immediate change from light color to dark color. Based on the order, the gradient is classified as positive and negative.  
**Positive Gradient:** color change from dark to light is called positive gradient.  
**Negative Gradient:** color change from dark to light is called negative gradient.  
**Horizontal Gradient:** moving from left to right gives us a horizontal gradient.  
**Vertical Gradient:** moving from top to bottom gives us a vertical gradient.  
Let’s look at how hog works, firstly the image which is in RGB is converted into a grayscale image because these images are perfect to work within the context of image processing. In the grayscale image consider the specified pixels and calculate horizontal and vertical gradients. Once we got the horizontal and vertical gradients it calculates two other parameters gradient magnitude and gradient angle. Based on these values the 64 gradient vectors are compressed to 9 vectors. This process is carried throughout the image and the results are combined to form our final whole image. The final image will result in only the content we required removing the rest of the useless data.
4.3 SVM:
Support vector machines (SVMs) comes under the category of supervised learning models that are used to interpret and analyze data for classification problems as well as regression problems. It solves different types of problems related to linear as well as nonlinear and is beneficial in a variety of contemporary applications. Implementing SVM is easy: This method draws a line or a hyperplane to partition data into classes or categories. SVM also works when we have more than two class instances as input (Multi Group classification). The retrieved HOG features are input into a set of Support Vector Machines. The goal of employing SVM is to get outcomes in a faster, more accurate, and simpler method. After extracting values for the input face image from HOG, those values which are calculated are measured with the values of the training set data using the Support Vector Machine Algorithm. After extracting facial features from the picture and then checks whether the driver in the picture is blinking his eyes for consecutive 15 frames or yawning mouth then the application will then send drowsiness signals to the driver. We are using SVM trained drowsiness model which is trained on a dataset that consists of four types of images like eye-opening, eye closing, yawning, not yawning and that will be used in our system it takes input from hog and mark the coordinates of eyes and mouth then we are continuously monitoring and predicting EYES and MOUTH aspect ratios, which will be determined using the Euclidean distance formula available in the SciPy package, based on which we will assess whether the driver is drowsy or not, the application would then warn the driver.

5. Results and Discussion
The System proposed in this paper is implemented using two classification algorithms, HOG and SVM. SVM is one critical algorithm used to detect feature which can be trained using a data set. SVM will also take a pre-trained model as an input. Some of the sample screenshots which shows how the proposed system will detect the drowsiness are given below, figure 3 shows how the application identifies the feature like eyes and mouth and draws a thin line around them and figure 4 shows how the application plots the 68 facial landmarks on the face of the driver/subject.
The application uses the facial landmarks that were shown above to not only to represent the feature with a thin line but also for calculation purpose. The aspect ratio calculations are done using the distances between this landmarks and drowsy alert is given accordingly. The figure 5 and figure 6 depicts the situations where the drowsiness alert is given to the driver i.e., when the driver closes his/her eyes for 15 consecutive frames or when the driver yawns.

6. Conclusions

Fatigueness of the driver may seriously affect driving abilities and endanger drivers and other car drivers. As discussed in related work, in present system there is poor performance in detecting the drowsiness of the driver. Therefore, we propose a better detection algorithm which considers multiple features from the face. In proposed system we are using two classification algorithms, a feature descriptor and a predictor. HOG is used as a feature descriptor which is used to identify object such as face of person. SVM is the other algorithm which is used as a predictor it takes input from HOG and then it detects eyes and mouth and judge them based on their size. Here the accuracy is improved and also the robustness of algorithms. So, we conclude that our suggested system has shown promising results and also outperformed others in terms of performance.
References

[1] Amir Jalilifard, Ednaldo Brigante Pizzolato"An efficient K-NN approach for automatic drowsiness detection using single-channel EEG recording”, 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

[2] T. Vesselenyi1, S. Moca1, A. Rusl1, T. Mitranl, B. OF Tătarul” Driver drowsiness detection using ANN image processing”1 2017 IOP Conf. Ser.: Mater. Sci. Eng. 252 012097.

[3] R. Savitha, T. Menaka, G. Thiruraman, M. Udhayan, “Real Time Driver Drowsiness Detection Using Convolutional Neural Network,” IJSART - Volume 6 Issue 3 – MARCH 2020

[4] J. Xu, W. Zeng, Y. Lan, J. Guo, and X. Cheng. “Modeling the parameter interactions in ranking SVM with low-rank approximation,” IEEE Trans. Knowl. Data Eng., vol. 31, no. 6, pp. 1181–1193, Jun. 2019.

[5] X. Kong, X. Liu, B. Jedari, M. Li, L. Wan, and F. Xia, “Mobile crowdsourcing in smart cities: Technologies, applications, and future challenges,” IEEE Internet Things J., vol. 6, no. 5, pp. 8095–8113, Oct. 2019. [39] A. Tolba, O. Said, and Z. Al-Makhadmeh, “MDS: Multi-level decision.

[6] Ashwani Kumar, Sonam Srivastava, “Object Detection System Based on Convolution Neural Networks Using Single Shot Multi-Box Detector”, Procedia Computer Science, Volume 171, 2020, Pages 2610-2617.

[7] Ashwani Kumar, S. S. S. Reddy and V. Kulkarni, "An Object Detection Technique For Blind People in Real-Time Using Deep Neural Network," 2019 Fifth International Conference on Image Information Processing (ICIIP), Shimla, India, 2019, pp. 292-297, doi: 10.1109/ICIIP47207.2019.8985965.

[8] Ashwani Kumar, "A Review on Implementation of Digital Image Watermarking Techniques Using LSB and DWT" in the Third International Conference on Information and Communication Technology for Sustainable Development (ICT4SD 2018), held during August 30-31, 2018 at Hotel Vivanta by Taj, GOA, INDIA.

[9] Sai Satyanarayana Reddy S., Kumar A. (2020) Edge Detection and Enhancement of Color Images Based on Bilateral Filtering Method Using K-Means Clustering Algorithm. In: Tuba M., Akashe S., Joshi A. (eds) ICT Systems and Sustainability. Advances in Intelligent Systems and Computing, vol 1077. Springer, Singapore.

[10] Kumar, Ashwani and Seelam Sai Satyanarayana Reddy, editors. Advancements in Security and Privacy Initiatives for Multimedia Images. IGI Global, 2021. http://doi:10.4018/978-1-7998-2795-5

[11] A. Tolba, O. Said, and Z. Al-Makhadmeh, “MDS: Multi-level decision system for patient behavior analysis based on wearable device information,” Comput. Commun., vol. 147, pp. 180–187, Nov. 2019.

[12] Z. Ning, Y. Feng, M. Collotta, X. Kong, X. Wang, L. Guo, X. Hu, and B. Hu, “Deep learning in edge of vehicles: Exploring trirelationship for data transmission,” IEEE Trans. Ind. Informat., vol. 15, no. 10, pp. 5737–5746, Oct. 2019.

[13] M. Khan and S. Lee, “A comprehensive survey of driving monitoring and assistance systems,” Sensors, vol. 19, no. 11, p. 2574, 2019

[14] Abbas, Q.; Alsheddy, A. Driver Fatigue Detection Systems Using Multi-Sensors, Smartphone, and Cloud-Based Computing Platforms: A Comparative Analysis. Sensors 2021, 21, 56. https://dx.doi.org/10.3390/s21010056

[15] Lee, J.; Kim, J.-W.; Lee, J. Mobile Personal Multi-Access Edge Computing Architecture Composed of Individual User Devices. Appl. Sci. 2020, 10, 4643