DRIVER's EYE: A Behavioural Monitoring System using Deep Learning

Sijo Joseph, S I Adithyan, Jishnuraj R, Aswin S
UG Scholar: Department of Computer Science
College of Engineering, Perumon
Kollam, India

Varun Chand H
Assistant Professor: Department of Computer Science
College of Engineering, Perumon
Kollam, India

Abstract— **Nowadays road accidents across the globe have increased a lot. The driver's distraction is one of the usual causes of such accidents. This results in health-related issues, loss of life, damages to vehicles, financial problems and sometimes it may create traffic congestions as well. The distractions can be reduced to an extent by using an alarming system when it detects the driver distraction. The paper proposed a system that intends to observe real time eye movement detection and in-vehicular signals. These observed data are fed into the model, which categories the driver movements as either distracted or normal. If distraction detects the alarming system is activated, which helps the driver to get back to the normal state. Viola Jones algorithms along with other machine learning algorithms are used for in-vehicular face analysis and vehicle signal analysis.**

Keywords—Viola Jones; Deep Learning; Driver Distraction; In-vehicular Signals; Face Data Analysis.

I. INTRODUCTION

For the past few years road accidents increased a lot, which resulted in serious health issues and even resulted in death [1, 2]. There are several risk factors associated with drivers, which finally results in accidents ranging from very minor to severe in nature [3]. The identified risk factors are professional driver, fatigue, large vehicle type, overload, and terrain. By providing technological solutions to these identified risk factors can help to reduce accidents up to an extent. With advancements in technology various smart and intelligent techniques are widely used in the transportation system [4-6]. Such techniques help to optimize road traffic, allocate parking space, etc. These techniques help to reduce driver frustration, recklessness etc. Different driver monitoring systems are currently available that focus on driver eye movement, behavior, fatigue etc.

The basic idea behind the project "Drivers eye" is real time monitoring of the driver's eye and facial features using a camera. Then these images are processed to find the drivers behavioural conditions like distraction, drowsiness, fatigue etc. The signals from the vehicle parts like accelerator, brake, steering etc. are also recorded and both these data are simultaneously processed to find out whether the driver is distracted. If the driver is distracted, the alarm system which is placed inside the vehicle will warn the driver. This will help the driver to recover from distracted state to the normal driving state. The majority of driver distractions are caused due the usage of smartphone, social networking activities etc.

while driving [7]. The proposed system can reduce the driver accidents that are caused due to such social activities.

The proposed model uses Viola Jones Algorithm for Face analysis and combined with other machine learning algorithms for evaluation of both In-Vehicular and Face data evaluation. The open data datasets were used for training both In-Vehicular and Face Data.

II. BACKGROUND ON DEEP LEARNING TECHNIQUES AND IN-VEHICULAR DATA

A. Viola Jones Algorithm

The Viola-Jones algorithm, created by Paul Viola and Michael Jones in 2001, is an object-recognition algorithm that enables for real-time identification of visual properties. Despite being an ancient algorithm, Viola-Jones is highly powerful, and its application in real-time face detection has proven to be exceptional. The Viola-Jones Algorithm has two stages:

      i. Training
      ii. Detection

i. Training

The algorithm reduces the image to 24 by 24 pixels and searches for the taught features. To be able to see features in their various and variable forms, it requires a large amount of facial image data. That is why, in order to train the algorithm, we must provide a large amount of facial image data. Viola and Jones fed 4,960 photos to their algorithm (each manually labeled). You can transmit the mirror image of a given image to a computer, which would be brand-new information.

You'll also need to give the algorithm non-facial images so it can tell the two classes apart. Viola and Jones provided 9,544 non-facial images to their system. Some of the photographs may appear to be features of a face, but the algorithm will recognise which features are more likely to be on a face and which features are clearly not.

1. Adaboost (Adaptive Boosting)

The algorithm learns from the photographs we provide it with and can identify false positives and genuine negatives in the data, making it more accurate. Once we've looked at all of the conceivable places and combinations of those traits, we'll have a highly accurate model. Because of all the possible possibilities and combinations you'd have to verify for each single frame or image, training might be quite extensive.

2. Cascading

Cascading is another "hack" we may use to improve our model's speed and accuracy. So we take a sub-window and look for our most important or finest feature within that sub-window to see if it is there in the image within the sub-window. We don't even look at the sub-window if it's not in the sub-window; we just dismiss it. If it's present, we'll look at the sub-second window's feature. If it isn't present, the sub-window is rejected. We keep going till we get to the amount of features, and then we reject the sub-windows that don't have the feature. Although evaluations may take only a few seconds, doing so for each characteristic can take a long time. This process is accelerated by cascading.

ii. Detection

Viola-Jones was created for frontal faces, therefore it can recognise frontal faces better than sideways, upwards, or downwards faces. The image is transformed to grayscale before detecting a face since it is easier to work with and there is less data to process. The Viola-Jones algorithm discovers the position on the coloured image after detecting the face on the gray scale image. Viola-Jones draws a box and searches within it for a face. It's basically looking for haar-like characteristics, which will be detailed later. After cycling through all of the tiles in the picture, the box moves a step to the right. A number of boxes identify face-like traits (Haar-like features) in smaller phases, and the data from all of those boxes combined aids the algorithm in determining where the face is.

1. Haar-like Features

Characteristics similar to those of Haar are named after Alfred Haar, a 19th-century Hungarian mathematician who invented the Haar wavelet concept (kind of like the ancestor of haar-like features). The features below depict a box with a light and dark side, which the machine uses to determine what the feature is. As in the edge of a brow, one side may be lighter than the other at times. The Center area of the box may be shinier than the surrounding boxes, which can be mistaken for a nose.
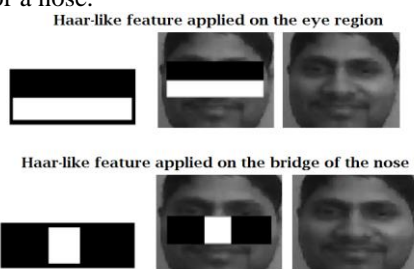

Fig 1 Haar like feature application
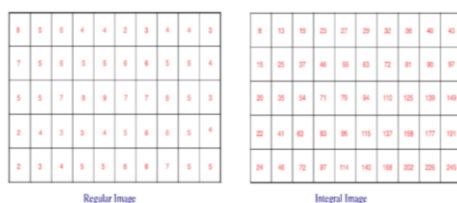
Integral Image

Fig 2 Integral images

The integral picture aids us in doing these time-consuming computations fast so that we may determine whether a feature of a group of features meets the criteria.

The total of the highlighted region in the regular image is used to determine the green box in the integral image. We'd have a sequence flowing through the grid if we did this for each box, and it might look like Fig 2.

2. Control Area Networks

A Controller Area Network (CAN bus) is a reliable vehicle bus standard that enables microcontrollers and devices to interact with each other's applications without the need for a host computer. It's a message-based protocol that was initially developed to conserve copper by multiplexing electrical wiring in vehicles, but it's now utilised in a variety of different applications. The data in a frame is broadcasted sequentially for each device, but in such a way that if many devices transmit at the same time, the highest priority device can continue while the others back off. All devices, including the sending device, get frames.

In 1985, Bosch created the Controller Area Network (CAN) for in-vehicle networks. Automotive manufacturers used point-to-point wiring methods to link electronic gadgets in automobiles in the past. Manufacturers began to incorporate more electronics into their automobiles, resulting in enormous wire harnesses that were both heavy and costly. They then used in-vehicle networks to replace dedicated wiring, which lowered wiring cost, complexity, and weight. The CAN protocol, which is a high-integrity serial bus system for networking intelligent devices, has become the industry standard in-vehicle network. CAN was immediately embraced by the automotive industry, and in 1993, it was designated as an international standard known as ISO 11898. CAN has formed the basis for numerous higher-level protocols, such as CAN open and Device Net, since 1994. These new protocols, which are now industry norms, have been widely adopted in other areas. Vehicle signals are passed using modern May networks, and we can readily access them from a CAN.

III. LITERATURE REVIEW

M. Sabet and colleagues introduced a unique technique for detecting driver tiredness and distraction [8]. A new module based on visual information and Artificial Intelligence is proposed here for autonomous driver sleepiness detection.

A. Warning System Design Face Detection

The suggested system employs the Viola and Jones (VJ) approach for face detection.

B. Object Tracking

They use the Ning et al. approach [9] for object tracking. The mean shift is used in this procedure. This method used the LBP pattern in addition to a colour histogram and proposed a joint color-texture histogram method, which represents a very distinct and effective target. The LBP operator is used to label pixels in a picture. This procedure was carried out by,
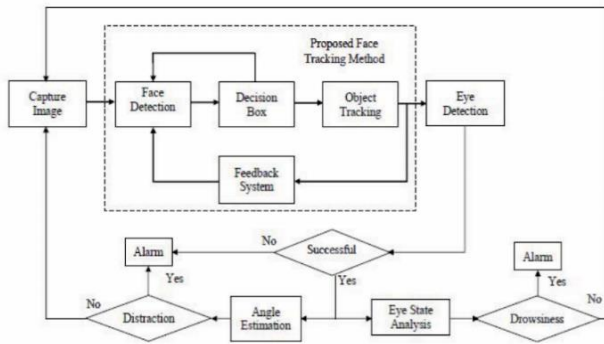
Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCREIS - 2021 Conference Proceedings

Fig.3 Block diagram

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p$$

The function s(x) is defined as follows: where g_c is the grey level of the centre pixel and g_p is the grey level of P pixels around the centre with radius R.

C. Feedback System

After the reference and output LBP pictures have been computed, the discrepancies between the two images can be measured using two-dimensional correlation. The lower the divergence, the higher the number.

When the divergence exceeds the threshold, the feedback system sends control to the detection system, which uses it to determine the true face position and re-initialize the object tracking system.

D. Decision Rules

The decision box always passes control to the detection system until the face is found at the start of the proposed system and when the first frame is loaded into the system. If the detection system fails to find the face after passing this level, the decision box sends control to the next section and then initialises the tracking system with the latest face location.

If the assumption of mild head movement is correct, and the face detection system outputs a location that is far from the previous position in each frame, we can assume the face detection system made a mistake and the output is a false alarm. As a result, the decision box transmits control to the next section, which utilises the face's most recent position.

E. Eye Detection

After obtaining the position of the face, it is possible to locate the eye with greater precision. Adaboost is a programme that detects the presence of eyeballs.

F. Eye State Analysis

The LBP operator was used to extract eye characteristics when the eye was discovered. Because LBP is not affected by light, eye state analysis can be more precise. The feature vector obtained from the sub-window is used in the next stage. As follows, compute feature vector d from LBP image G(x, y) with sub-window size 5 by 6:.

G. Distraction

The driver's face should be investigated to detect this trait because the attitude of the face carries information about one's attentiveness, gaze, and level of exhaustion. The following approach has been implemented to verify driver distraction. The following equation is used to calculate face orientation using the eye position:

In this study, a new approach for monitoring and detecting driver tiredness and distraction is provided. This system makes use of cutting-edge computer vision and artificial intelligence technology. To avoid divergence and losing the target, we used a feedback system based on LBP for face tracking. The detection and tracking modules can work together thanks to the feedback mechanism.

SVM was used to analyse the eye condition, and the LBP operator was used to extract features. For variable light, background changes, and facial orientation, the suggested method for face tracking and eye state analysis proves to be resilient and accurate. It's also been noted that the system generates agreeable outcomes.

S. Maralappanavar and colleagues suggested a method for detecting driver distraction using gaze estimation to estimate the driver's attention and evaluate whether the vehicle is distracted or not [10]. The driver's look direction is used to gauge his attentiveness. The driver's gaze is estimated by detecting the look with the help of the face, eye, pupil, and eye corners, and then categorising the discovered gaze as distracted or not.
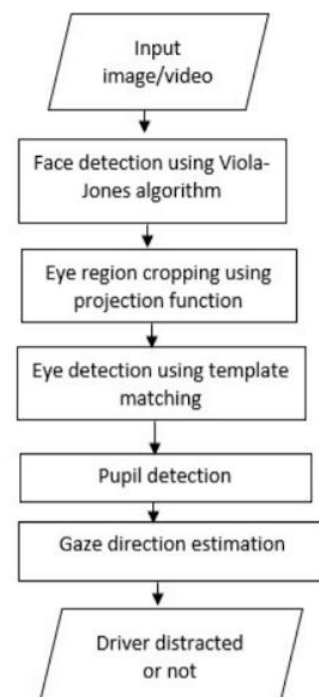

Fig. 4 Proposed Algorithm for Driver's Eye Detection Systems

Different modules are used to implement the drivers' gaze direction estimation. Face detection is done first, followed by accurate eye detection using the template matching method, and then gaze estimation. The driver's gaze is used to determine how distracted he or she is. Face Detection Using

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCREIS - 2021 Conference Proceedings

Viola-Jones Approach is used a flowchart of the suggested algorithm is given

H. Eye Detection with Template Matching Method

The following is how the template matching is done: An eye template of size w x h is created, and an input image of size W x H is created. By overlapping the template picture on the input image, the normalized cross-correlation is discovered. The size of the input image is adjusted to fit the template. In our situation, three templates are used: one for the left, one for the right, and one for the centre. The maximum value for the matched place is returned by template matched. The highest value for each of the templates is compared to the highest value obtained by combining the three. The white box indicating the template matched output surrounds the matched places. The eye region is accurate thanks to the template-matched output.

I. Gaze Direction Estimation

Horizontal distance is calculated using the distance between the pupil boundary and the eye corners. On both eye corners, a line segment is drawn. The distance between one pupil boundary and one eye corner line segment is then calculated as d, while the distance between the other pupil boundary and the other eye corner is calculated as d1. The position of the pupil and the placement of the eye corner are utilised to determine the direction in which the person is gazing. The calculated distance is used to determine whether or not the driver is distracted. The eye corner segments and pupil boundaries are depicted in Fig. 5..
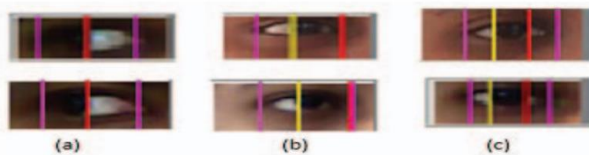

Fig. 5 Segmentation of Eye

Face detection is performed using the viola jones algorithm, followed by eye region localization using the integral projection function, pupil detection, and gaze categorization in their suggested system.

L. Alam and colleagues proposed a real-time distraction detection system [11] based on the driver's visual characteristics. Its goal is to create a method for detecting distraction in real time by assessing a driver's visual features in the facial region.

Fig 6. Schematic diagram of driver's attention monitoring system

The suggested method extracts essential information from visual cues such as eye and head movement to detect driver attention states and classify them as either attentive or distracted. In this strategy, the deviation of the eye centre and head from their normal positions over a period of time is considered to be useful signs for identifying lack of attention.

Face detection is conducted first, followed by the extraction of the region of interest (ROI) - the eye and head region - using facial landmarks, and finally the detection of head and eye movements to identify the attention state. They conducted an experiment in a real-world driving setting with people with various characteristics to assess the system's performance. In all of the cases they evaluated, their system detected the attention state with an average of 92 percent accuracy.

J. Eye Movements Detection

The centre of the eyes is detected to estimate the rate of departure of eye centres from the standard position in order to identify eye movement. The location of the eye's centre has been discovered in order to do so at first. Some preprocessing on both the left and right eye pictures found by six (x, y)-coordinates from the preceding modules is performed individually to facilitate the determination of the eye centre. The sites relating to the ocular region that were used to identify eye movements are shown in Fig. 7.
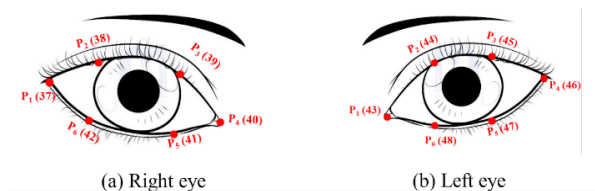

(a) Right eye   (b) Left eye

Fig 7. Eye Region Detection

The video stream taken from a web cam-era installed on the dashboard is used in this work to develop a system that detects distraction in real-time. The system may detect distraction in the driver based on a range of face features (such as facial features, haircuts, and accessories). For the entire situation, the system had an accuracy of roughly 92 percent. Our tests revealed that our system for detecting eye and head movement is reliable. The addition of a warning system to

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCREIS - 2021 Conference Proceedings

alert the driver if distraction is detected is being maintained as a future project.

No distraction detection by in-vehicle signal processing was carried out by S. Im and colleagues [12]. The techniques and experimental findings for detecting distraction without deliberate distraction using in-vehicle signals. Normal and distracted driving features can be classified in real-world driving circumstances by combining two types of machine learning algorithms - unsupervised learning and supervised learning.

K. Signal Selection and Data Processing

Many in-vehicle signals are available, including steering angle, steering speed, acceleration/brake pedal value, yaw rate, longitudinal/lateral acceleration, and so on. Finding signals relating to driving style is critical. These signals should also have characteristics that distinguish between normal and distracted states. They discovered signals that met these criteria. These signals were chosen and analysed in order to derive characteristics of both regular and distracted driving.

L. Data Processing

Before being used as inputs to the learning system, selected signals should be preprocessed. First and foremost, they employed signals that indicate a vehicle speed of more than 60 km/h. They assumed, like other automakers' products, that the algorithm works on roads. When the vehicle speed is modest, it's difficult to tell the difference between regular and distracted driving.
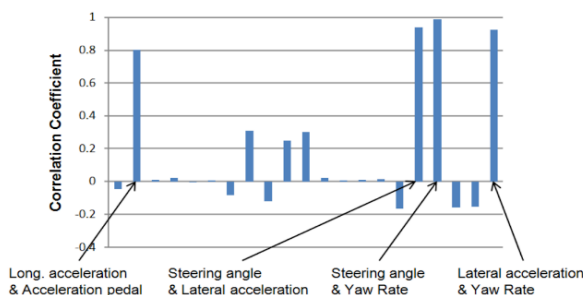

Fig 8. Correlation coefficient of in-vehicle signal combinations

	Average		CV		Standard Deviation	
	Dist.	Norm.	Dist.	Norm.	Dist.	Norm.
Acceleration pedal value	10.3	12.1	1.03	0.94	10.7	11.4
Brake master cylinder pressure	0.60	0.30	3.56	3.98	2.15	1.20
Relative distance to front vehicle	55.3	66.7	0.81	0.67	44.6	44.8
Relative speed to front vehicle	-0.13	-0.34	-13.9	-5.85	1.75	2.01
Distance from a wheel to a lane	2.73	2.81	0.23	0.21	0.63	0.58
Steering wheel torque	0.09	0.08	12.0	13.1	1.02	1.05
Steering wheel angle	0.27	0.24	39.4	10.2	10.8	2.45
Steering wheel speed	1.66	0.81	4.11	2.94	6.81	2.38

Table 1. Statistical Analysis

M. Distraction Detection Algorithm Using Machine Learning

In terms of learning strategy, the driver distraction detection algorithm works in two steps: first, distraction detection using only normal driving patterns, then, if specific conditions are met, distraction detection using both normal and distracted driving patterns.

Because it uses only normal driving data to generate the model, the distraction detection strategy employing the GMM model cannot create the decision border between normal and distracted driving. In most experiments, supervised learning systems were utilised, which required not just regular driving data but also inattentive driving tendencies. Their study likewise used supervised learning, but the distracted driving pattern was obtained in a different way. They did not need drivers to perform any additional duties, but they did see a distracted driving tendency. The statistical disparity between normal driving data and test data is high when drivers are preoccupied while driving. To get a more reliable pattern, more thresholds of statistical distance are needed.

They devised an algorithm that combines two approaches and learning schemes: normal/distracted driving pattern learning (GMM) and normal/distracted driving pattern learning (NDPL) (ANN). They demonstrated that distracted driving data can be collected naturally on real roads. The algorithms' two phases were compared, and it was discovered that the way that uses two types of driving data performed better. They have no clue about the state of the driver other than how he employs an accelerator, brake, and so on, thus not only quantitative but also qualitative evaluation is vital.
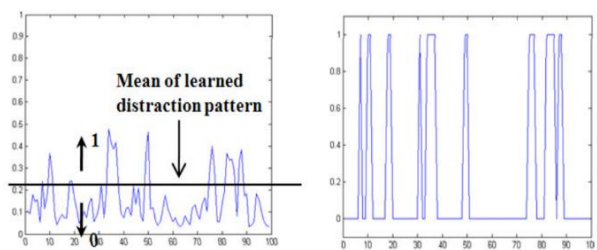

Fig 9. An example of ANN output

As a new technique to detecting distracted car drivers, F. Mizoguchi and et al. construct new rules for determining whether or not a driver is distracted, using acquired data about the driver's eye movement and driving data through learning [13]. To generate the rules, they employ a learning technology called a support vector machine (SVM). In addition, they explored the relationship between a qualitative model of a driver's cognitive mental load and the driver's distraction in a previous study. We validate the driver's eye movements and driving data that are contradictory with the model during the investigation.

N. SVM learning

Figure 10 shows data for the case where the SVM learns the collected parameters. One saccade corresponds to the parameters on a single line. The number on the left (+1 or -1) indicates whether you are not distracted (+1) or distracted (-1). (-1). After scaling, the SVM can learn these parameters, and learning data is generated. For SVM learning, they used the Gaussian kernel as a kernel function.

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCREIS - 2021 Conference Proceedings

Fig 10. Parameters used in SVM learning

Through a new approach to detect distracted driving that incorporates an SVM learning tool, their study creates new rules for visually dis-tracted driving utilising acquired data on the driver's eye movement as well as driving data. They can ensure that precision can be reached when driving on a road with moderate traffic by confirming the rules established via the learning process.

Furthermore, they analyse the relationship between a qualitative model of a driver's cognitive mental load pro-duced in a previous study and the driver's distraction. They look for any eye movement or driving data that contradicts the model during the examination. As a result, they can detect visually distracted driving and confirm that the frequency of saccades per unit time during driving has decreased. This behaviour is in line with the characteristics of cognitive dis-traction. It is possible to detect distraction during operation by noting the mental-load qualitative model, which reveals that there are ways beyond interpreting the driving condi-tions.

IV. PPROPOSED SYSTEMS

Proposed methodology is to introduce an alarming system that alert driver when he is distracted, we use the eye move-ment pattern analysis data along with in vehicular signal pat-tern analysis data to evaluate the same, our system gets trained with collected eye and in vehicular signal patterns which then can evaluate real time patterns of both.

The proposed model uses viola jones algorithm for face data analysis, then combines it with in-vehicular signal analy-sis data using a new algorithm

By using both data the model can form two different trained models then the results of in vehicular data model can be used to decide the thresholds for face data analysis. Eg. When we see a very low speed(<20kmph) we can allow the driver to look away for a little more time than usual.
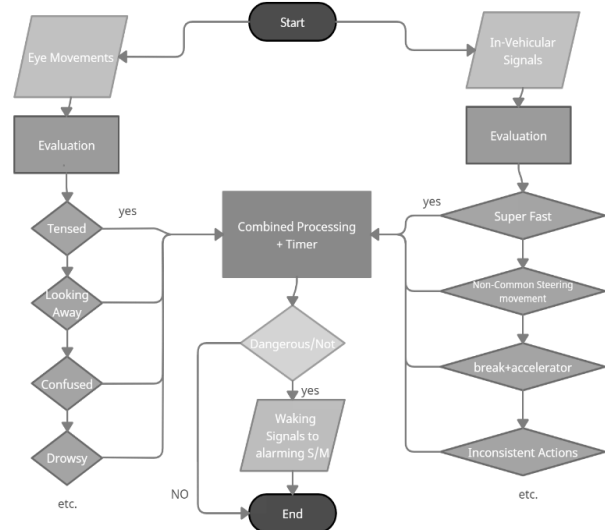
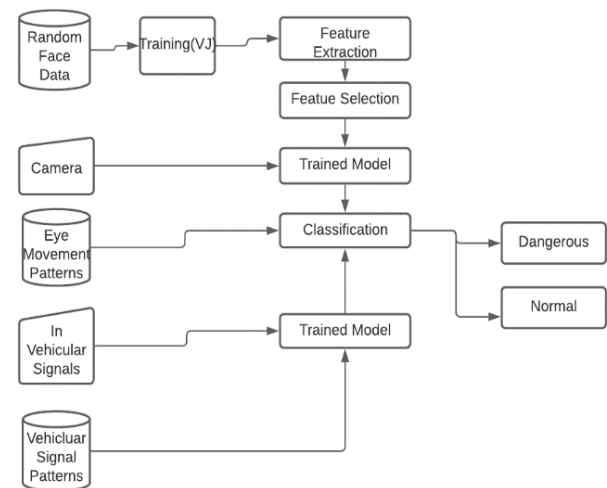

Fig 11. Flow Chart for DRIVER's Eye Algorithm
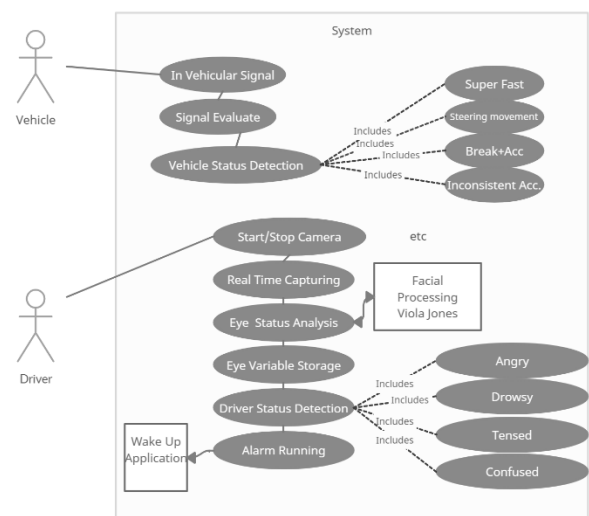

Fig 12. DFD Diagram

Fig 13. Use Case Diagram

We can even make use of open source pre-trained models that are already good enough for face data evaluation. E.g. shape_predictor_68_face_landmarks (Dlib).

For the In-vehicular data analysis model described in [12] some additions with respect to Indian roads can be done. Both data are evaluated with their own haar like features then o/p of both models took together to decide the thresholds for facial feature monitoring.

V. CONCLUSION AND FUTURE STUDY

A driver eye monitoring system has been proposed in this paper to alert the driver when his face is distracted from the normal state. Real time analysis of driver's eye movements and in vehicular signals like signals from break, accelerator, steering etc. helps in deciding whether a driver is in a distracted state or in a normal state. Both eye movement and vehicular signal data are used for converging a better solution for identifying distraction in the driver. The proposed model can be built with driver's eye movements as input from the camera and the state of driver can be evaluated using a vehicular system (CAN). The alarming system will receive a signal when the driver's state changes to distraction and produces a sound to alert the driver. By this way the driver can promptly return to the normal state from the distracted one.

Additional features can be added to improve the performance of this proposed model. If the cause of distraction is uncommon like, distracted by seeing an accident, misbehavior from co-passenger, road problems etc. the driver can inform such things to concerned authorities. This can be implemented by adding add-on features to the proposed model of EYE monitoring and alerting system. The proposed model can even check whether the driver is minor or not. If the driver is a minor such information can be passed to the concerned authority.

The proposed system combines two of the most common methods and uses more refined algorithms. Hence the model is expected to produce precise and faster results.

REFERENCES

[1]. S. K. Singh, "Road Traffic Accidents in India: Issues and Challenges,"*Transportation Research Procedia,* vol. 25, pp. 4708-4719,2017.

[2]. Y. Darma, M. R. Karim and S. Abdullah "An analysis of Malaysia road traffic death distribution by road environment," *Sadhana*, vol. 42, pp. 1605–1615, 2017.

[3]. G. Liu , S. Chen, Z. Zeng, H. Cui, Y. Fang and et.al, "Risk factors for extremely serious road accidents: Results from national Road Accident Statistical Annual Report of China," *PLoS ONE*, vol. 13, no. 8, 2018.

[4]. H. V. Chand, J. Karthikeyan, "Recent Survey on Traffic Optimization in Intelligent Transportation System," *Journal of Advanced Research in Dynamical and Control Systems,* vol. 9, no.15 ,pp. 89-100, 2017.

[5]. J. Zhang, Y. Yu and Y. Lei, "The study on an optimized model of traffic congestion problem caused by traffic accidents," *2016 Chinese Control and Decision Conference (CCDC)*, pp. 688-692, 2016.

[6]. E. C. Eze, S. Zhang, E. Liu and J. C. Eze, "Advances in vehicular ad-hoc networks (VANETs): Challenges and road-map for future development," *International Journal of Automation and Computing,* vol. 13, pp. 1-18, 2016.

[7]. F. D'Amico, A. Calvi, C. Ferrante and L. B. Ciampoli, "Assessment of Driver Distraction Caused by Social Networking Activities Using the Smartphone: A Driving Simulator Study," In: Stanton N. (eds) *Advances in Human Aspects of Transportation,* Advances in Intelligent Systems and Computing, vol. 1212, Springer, Cham, 2020.

[8]. M. Sabet, R. A. Zoroofi, K. Sadeghniiat-Haghighi and M. Sabbaghi-an,"A New System for Driver Drowsiness and Distraction Detection," *20th Iranian Conference on Electrical Engineering (ICEE2012)*, pp. 1247-1251, 2012.

[9]. J. Ning, L. Zhang, D. Zhang and C. Wu, "Robust mean-shift tracking with corrected background-weighted histogram," *IET Computer Vision,* vol. 6, no. 1, pp. 62-69, 2012.

[10]. S. Maralappanavar, R. Behera and U. Mudenagudi, "Driver's Distraction Detection based on Gaze Estimation," *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2489-2494, 2016.

[11]. L. Alam and M. M. Hoque, "Real-Time Distraction Detection Based on Driver's Visual Features," *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 1-6, 2019.

[12]. S. Im, C. Lee ,S. Yang, J. Kim and B. You, "Driver Distraction Detection By In-Vehicle Signal Processing,*" 2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*, pp. 64-68, 2014.

[13]. F. Mizoguch, H. Nishiyama and H. Iwasaki, "A New Approach to Detecting Distracted Car Drivers Using Eye-Movement Data," *2014 IEEE 13th International Conference on Cognitive Informatics and Cognitive Computing,* pp. 266-272, 2014.