www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 3.1, (2026), pp. 360-368
Open Access. Online & Print

DOI: 10.5281/zenodo.21469375

X-DRIVE: AN EXPLAINABLE AND RESPONSIBLE AI FRAMEWORK FOR REAL-TIME DRIVER ATTENTION AND DROWSINESS MONITORING

Amol Raoji Madane[1*], Ho Chin Kuan[2]

[1]Post-Doc fellow, Asia Pacific University of Technology, Jalan Technologi 5, Taman Technologi, Malaysia, 57000 amol.madane@gmail.com
[2]Vice Chancellor and Post-Doc Research Guide, Asia Pacific University of Technology, Jalan Technologi 5, Taman Technologi Malaysia, 57000 ckho@apu.edu.my

Received: 12/03/2026
Accepted: 03/04/2026

Corresponding Author: Amol Raoji Madane (amol.madane@gmail.com)

ABSTRACT

Driver distraction and fatigue remain among the leading causes of road accidents worldwide, necessitating the development of intelligent, reliable, and human-centric driver monitoring systems. This paper presents a vision-based driver state monitoring framework for the real-time detection of inattentiveness and drowsiness using multiple behavioral indicators, including eye closure, yawning, head lowering, and head-pose-based attention assessment. Unlike conventional approaches that rely on a single distraction cue, the proposed framework integrates complementary visual features to improve detection robustness and reduce false alarms under varying driving conditions. The system employs computationally efficient facial feature extraction and classification techniques to continuously analyze driver behavior using an in-cabin camera without requiring intrusive physiological sensors. Furthermore, the framework utilizes interpretable behavioral indicators that enable transparent decision-making, supporting the principles of Explainable Artificial Intelligence (XAI) in safety-critical automotive applications. Experimental evaluation conducted on a diverse set of subjects demonstrated high detection accuracies across multiple distraction and fatigue scenarios, validating the effectiveness of the proposed approach for real-time deployment. The findings indicate that integrating multiple observable behavioral cues enhances driver state recognition and contributes toward the development of trustworthy, responsible, and intelligent Advanced Driver Assistance Systems (ADAS). Future enhancements include explainability-driven decision visualization, fairness assessment across diverse driver demographics, and multimodal sensor fusion for improved robustness and reliability in real-world environments.

KEYWORDS: Advanced Driver Assistance Systems (ADAS), Computer Vision, Driver Distraction Detection, Driver Monitoring System, Drowsiness Detection, Explainable Artificial Intelligence (XAI), Human-Centered AI, Responsible AI, Road Safety, Trustworthy AI.

1. INTRODUCTION

Advanced Driver Assistance System (ADAS) are vehicle-based rational safety systems. ADAS provides real time testing approach which alerts the driver about potential problems well in advance. ADAS monitors the driving environment and executes some driving tasks automatically. ADAS features, hence, assists the driver and provides necessary alerts about potential problems beforehand by analyzing the driving environment.ADAS systems provide drivers with enhanced safety features such as adaptive cruise control, automatic emergency breaking, guidance to change the lane and provide forward collision warnings. Other ADAS features include lane departure warning, automatic parking, night vision monitoring, drowsiness monitoring and traffic sign recognition. It implements multiple applications using data from radars, LiDAR, ultrasonic and infrared sensors.ADAS plays a major role in aiding distracted drivers. Driver distraction is considered as major factor for fatalities on roadways. Driver distraction occurs when a driver's attention is diverted towards other in-car activities. Distracted driving activities include adjusting vehicle setting, texting or having conversations with co-passengers, etc. which have serious consequences for road safety. Drowsiness is failure to perform the driving task due to lack of sleep, shift work or medications or driving at times of the day when you would normally be sleeping. Drowsiness affects driver's abilities of reaction, information processing and making judgment. It is very helpful to remind the drivers to take a rest or improve vigilance when they feel drowsy. The state of eyes reflects if the driver is drowsy or not. But that is not the only information which reflects driver's drowsy condition. Yawning and head lowering are other important factors that reflect driver's fatigue.

Inattentive driving is an inability to pay attention to the road while driving. It includes turning back to talking with co-passengers, eating, and using cell phone while driving. Inattention is categorized as Manual, Visual and Cognitive. Manual distraction is taking driver's hands away the wheel, while visual distraction occurs when driver's eyes wander off the road and cognitive distraction is when driver's mind is off from driving.As driver distraction has become major factor in road crashes, three approaches exist for driver distraction detection. They are methods based on driver's biological measures, driving performance measures and driver's behavioral measures. Driver biological measures include brain wave response, Heart Rate Variability (HRV), body temperature, etc. to estimate driver distraction. These techniques require physical contact with drivers (e.g. attaching electrodes) to perform detection and hence, are intrusive, causing annoyance to drivers. Second approach considers standard deviation of steering wheel position, mean of steering error or standard deviation of lane position. These techniques though less intrusive, but they are subjected to several limitations including the type of vehicle, driver's expertise in driving, driving conditions and external environment. The third approach which is driver behavioral measures examines the driver's physical condition such as eye gaze direction, eye closure rate, blink detection or head position. Our system comes under the category of driver behavioral measure which proposes the face monitoring approach focusing straight on driver behavior using camera to recognize distraction. The technique proposed in the paper is not intrusive and does not possess limitations due to driving environment or type of automobile. This is because an in-car camera is utilized to take video frames of driver and those video frames are processed using various algorithms proposed in the upcoming sections.

2. LITERATURE SURVEY

Many efforts have been reported in the literature on developing real-time distraction monitoring systems [1-5]. Inattentiveness detection is usually based on gaze movement [3] and head rotation [2,3]. Drowsiness detection is based on eye blinking rate [1,2,4,5] and yawning detection [1,2,5].Manu B.N [1] presents the real time method for drowsiness detection by three defined phases. It includes facial features detection using Viola Jones, eye tracking and yawning detection. Skin part is segmented alone after the face detection. Chromatic components are considered to reject the non-face image backgrounds based on skin color. Yawning detection and tracking of eyes are done using correlation coefficient template matching. The consecutive frames are classified into fatigue and non-fatigue states using binary linear support vector machine classifier. System detection fails on tilting of head towards left or right. Also, the head lowering state prediction has not been considered.

Yogesh et al [2] have proposed research in which physiological as well as physical signs are considered for detection of drowsiness. Physiological factors considered are body temperature and pulse rate. These factors will decrease during the onset of drowsiness. Physical visual cues considered are yawning, closed eyes, drooping eyelids and increased blink durations. Combination of these factors are used to detect drowsiness and alert the driver. In the technique mentioned, the driver needs to maintain physical contact with the sensors which become intrusive to the driver.Meng-Che et al [3] have proposed a method of driver gaze direction estimation technique in which many classifiers such as ccascade, SVM classifiers are used and new data collection method called in-situ is introduced. Inspite of improvement in generalization performance of classifier, there is demand of more efficient machine learning approaches for gaze estimation. H. Eren et al [4] proposed the system using inertial (smart phone) sensors, namely the accelerometer, gyroscope, magnetometer. Using deflection angle information, speed, acceleration and deceleration, the system statistically analyzes driver behavior. Driver behavior is estimated as safe or unsafe using Bayesian classification, and optimal path detection. M. Omidyeganeh et al [5] proposed fusion based system for driver drowsiness detection by considering eye closure and yawning detection methods. YCbCr and HSV color space model followed by morphological operations are used for face detection. Determination of eye closure is based on horizontal projection [7]. YCbCr color space is used for calculating the mouth map and morphological operators are added to it. Yawning detection is confirmed by continuous checking the hole in mouth map. Though system does not require any training algorithm, but accuracy of morphological operators suffer in noisy conditions and driver distraction detection is a real time based system which demands the processing of noisy image frames.

3. PROBLEM STATEMENT

We present a method for monitoring driver distraction by focusing on not only single visual signs of distraction but all the possible visual cues. We have two broad categories to define visual cases while driving:

- Safe = {“looking towards road”, “viewing left mirror”, “viewing right mirror”, “viewing top mirror”, “eyes are open”, “eyes are blinking”, “mouth is closed” “head is straight up”}.
- Unsafe = {“talking to co-passengers”, “looking outside the window, “eyes are closed”, “yawning”, “head is lowered”}.

System detects all safe and unsafe driver behavior and raises audio alarm for unsafe driving cases.

4. PROPOSED ALGORITHM

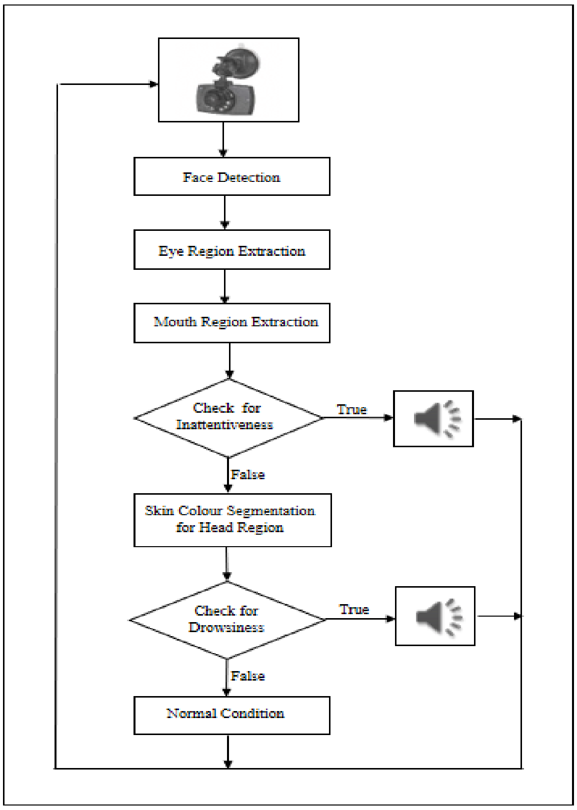


Figure 1: Flowchart of Proposed System.

System proposes separate approaches for detecting inattention and drowsiness. From the flowchart Fig.1, it is evident that the system is a continuous driver condition monitoring real-time system. Images given as input to the system are obtained from the in-car infrared camera installed at A-pillar of the car to get clear images of driver without interrupting forward view of driver. Images are processed frame by frame for detection of inattention followed by drowsiness [8]. Once the system detects inattention, an audio alert is raised and the next frame is inputted for processing. If the system detects the frame to be attentive frame, then it is checked for drowsiness. In Drowsiness detection, firstly, eye status is checked, followed by yawning and last is head lowering. Detection of any one of the three or combination of any two or all three confirms drowsiness. Audio alert is generated and next frame is inputted for processing.

4.1. Face Detection

The Viola Jones algorithm [6] is used for detection of the face from an input image. This approach combines three key concepts for the implementation of face detection algorithm. The first step includes calculation of the features which is facilitated with the use of an "Integral Image". It provides fast feature evaluation and efficient means to speed up the classification function of the system. The second step employs Adaboost to select the key features and to train classifiers. Adaboost algorithm takes a training set of positive (face) and negative (non-face) images to train classifier. Next step involves construction of a "strong" classifier as a linear combination of weighted simple "weak" classifiers. All the features are grouped into several stages where each stage has certain number of features. Each stage determines whether a given sub window is definitely a face or not a face. A given sub window is immediately discarded as not a face if it fails in any of the stage.

4.2. Eye Region Extraction

Face detection algorithm determines the location and size of face in an input image which is enclosed in a rectangular bounding box. It contains a four-element vector, [x y width height], that specifies in pixels, the upper left corner and size of a bounding box. A novel and robust eye region extraction algorithm is employed using symmetry-based approach of face. To reduce the computation time of algorithm, eye region extraction area is restricted to top half segment of facial image. In this technique [9], two individual regions (left and right eye region) containing eyes and eyebrows are separated from the region of total face. Each eye position is detected independently within an area divided by the centerline of the face, which is calculated from the face width as shown in Fig.2. Proposed algorithm utilizes simple and efficient image processing method which reduces the computational complexity. It gives remarkable results without any need of training.

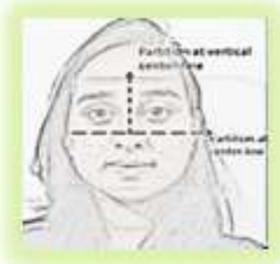


Figure 2: Eye Region Extraction.

4.3. Inattentiveness

Head pose provides information about the visual focus of attention. Hence driver's attention is determined by the head pose which is expected to be in front direction. In the proposed method, input video frames are first processed through frontal profile face detector to estimate head pose direction. Frontal profile face detectors use Adaboost cascade classifiers operating on Haar-like features. Adaboost cascade classifiers [6] is trained specifically to locate respective facial regions such as left eye, right eye, mouth and nose. Normalized face part locations and sizes are determined using 14-dimension feature vector as shown :

$$x = \left(x_{le}, y_{le}, s_{le}, x_{re}, y_{re}, s_{re}, x_n, y_n, w_n, h_n, x_m, y_m, w_m, h_m\right). \quad (1)$$

In Eq.1, *x* and *y* represent spatial locations, *s* stand for side length, *w* and h denotes width and height, respectively. Subscripts *le, re, n, m* stands for left eye, right eye, nose, and mouth respectively.

Frontal face is detected when all mentioned features are present in the frame. Non-frontal pose is encountered if any of these features is absent. Driver looking in the side mirrors of the vehicle is also a case of non-frontal pose. There is a need to distinguish between the case of driver is looking at side mirror and the case of inattention. Hence time-based approach is considered as distinguishable parameter. In the proposed algorithm, level of inattention is determined by consecutive frames of non-frontal profile. After encountering first non-frontal pose frame, if consecutive non-frontal frames are not obtained for more than 3 seconds, then this is considered as a case of driver looking in side mirror.

Whereas, inattention is confirmed when non-frontal profile frames are obtained for more than 3 seconds continuously.

4.4. Drowsiness

We are proposing a robust drowsiness detection system based on three parameters which are eye state, yawning and head lowering. Since the system does not focus on only single visual cue of drowsiness but all the possible visual cues, this makes detection invariant and works with greatest efficiency regardless of specific driver and operating conditions.

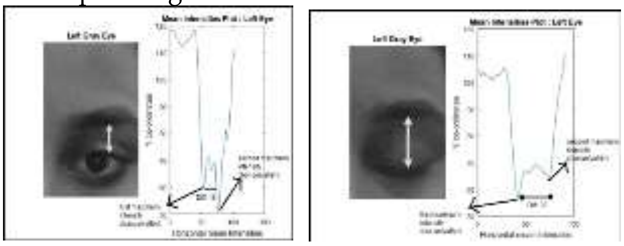
4.4.1. Drowsiness Detection Based on Eye State

Eye closure is the major indicator of driver fatigue. Our proposed algorithm detects the eye state of driver which is either 'open' or 'close' and processes the closed eye frame to detect drowsiness. Drowsiness is a process of continuous

eye closure whereas, eye blinking is a process of temporary eye closure. The system efficiently distinguishes eye blinking, and drowsiness and no false alarm is raised for eye-blinking [10].

4.4.2. Proposed Algorithm for Eye State Classification

The decision about the state of eye is based on the distances between eyebrows and upper edge of eyes. Fig.3. shows that the distance between the eyebrow and upper edge of eye for open eye is much smaller than the distance in case of close eye. These distances are calculated by plotting horizontal intensity change plot for both extracted left and right eye regions. Horizontal intensity plot is the plot of average of all horizontal pixels at y-coordinates. The horizontal plot has valley point which results from intensity changes along y-axis. The most significant valleys (lowest valleys) occurs across the eye and eyebrows since they are darkest (blackest) parts of the extracted eye regions. Hence, to calculate the distance between eyebrows and eyes, the differences in these two significant valleys are calculated. As the distances are smaller for open eyes than close eye, so is the difference between the valleys as shown in Fig.3(a). The decision about the state of eye is taken after calculating the valley differences [11]. The first frame on the start of the system is considered as the reference frame under the inference that when the system was started driver was attentive and awake. We consider the valley difference for open eye of the that particular driver seating at driving seat for whom the system is currently testing because the valley difference cannot be an absolute generalized value. This is because eyelid distances differ for different individuals across the globe. For example, distances in Japanese or Chinese men/women are lower than European, Mid-East or African men/women. So, to make the system robust and real time, the valley difference in the first frame is marked as open eye valley differences and stored to compare with upcoming frames.



(a) Open eye *(b) Close eye*

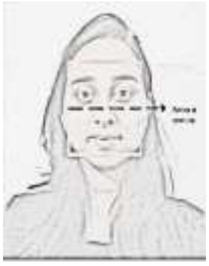
Figure 3: Distances between Eyebrows and the Upper Edge of the Eye Along with the Mean Intensity Plot.

By comparing the upcoming frames valley differences with reference values, the state of eye is classified. The closed eye frames are processed, and series of consecutive closed eye frames indicates drowsiness. By considering series of consecutive frames, system avoids eye blinking conditions and prevents false alarming. Therefore, after encountering first closed eye frame, consecutive frames are checked. When consecutive frames are not detected as closed eye for more than 3 seconds, it is considered eye blink. However, if closed eye frames are obtained continuously for more than 3 seconds, drowsiness is confirmed.

4.4.3. Yawning Detection

4.4.3.1. Mouth Region Detection

First step in mouth region extraction is obtaining the lower half-segment of facial image obtained in face detection section. It is obtained by partitioning at the horizontal centerline of the face, which is calculated from the face height as shown in Fig.4(a). Mouth search window is shown in Fig.4(b), is restricted to extracted lower half region of the face. It is done to limit mouth search to only lower area so as to reduce the computational complexity. The algorithm used for detection of mouth in the search window is based on Viola Jones. In this approach of mouth detection, the detection of the features in the mouth region is based on a decision stump which uses Haar features to encode the details of the mouth.



(a) Partitioning of the face.

(b)Mouth search window.
Figure 4: Mouth Region Extraction.

4.4.3.2. Proposed Algorithm for Yawning Detection

Yawning is a process of continuous mouth openness in which the distance between upper and lower lips is considerably larger as shown in Fig.5(a). Whereas, while in the case of talking this distance is comparatively smaller and there is temporary mouth openness as shown in Fig.5(b). The mouth detection is based on Viola Jones which is trained using Adaboost classifier to detect the closed mouth frames [12]. So, when the yawning or talking like frames (open mouth frames) appear for processing, the algorithm increases the merge threshold value and test for presence of closed mouth. But, these frames with open mouth could not qualify for mouth detection because our algorithm has predefined merge threshold value set to 200. This large threshold value ensures that only closed mouth and talking-like mouth frames qualify for mouth detection. In this way, we distinguish between the yawning and talking cases and detect the frames in which mouth openness is very large. Once the first frame of yawning is detected, the system checks for consecutive frame to confirm yawning so that algorithms do not detect talking condition to be yawning and raise a false alarm.

(a) Yawning.
(b) Talking.
Figure 5: Yawning Detection.

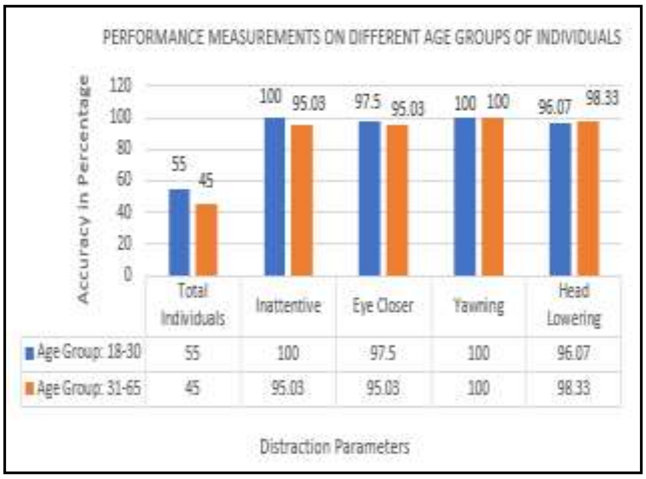
4.4.5. Head Lowering

Head orientation of the driver is also one of the major visual cue which provide information about driver's vigilance. When the driver feels sleepy, he/she experiences trouble keeping his/her head up. Hence, the system proposes head lowering detection algorithm to detect the drowsy driver. In this algorithm, colour space conversion technique is used. By colour space conversion, the RGB facial image is converted to YCbCr facial image. Ideal driving condition means driver's head direction is front-facing and straight up as shown in Fig.6(a). Skin segmentation is applied on the reference frame (first ideal frame on the start of system) and percentage of skin pixels of this frame is calculated and stored, Fig.6(b). Skin segmentation is applied on all the upcoming frames and skin pixels percentage is calculated. The comparison is made between the percentages of skin pixels in the current frame with the reference frame. Predefined threshold value is set. If the difference between the percentage of skin pixels in current and reference frame is greater than the threshold, the system detects the frame to be head lowered frame, as shown in Fig.6(c).

a) Frontal head direction (b) Skin segmented image (c) Head lowered image: No. of skin pixels less.
Figure 6: Head Lowering.

5. EXPERIMENTAL RESULTS

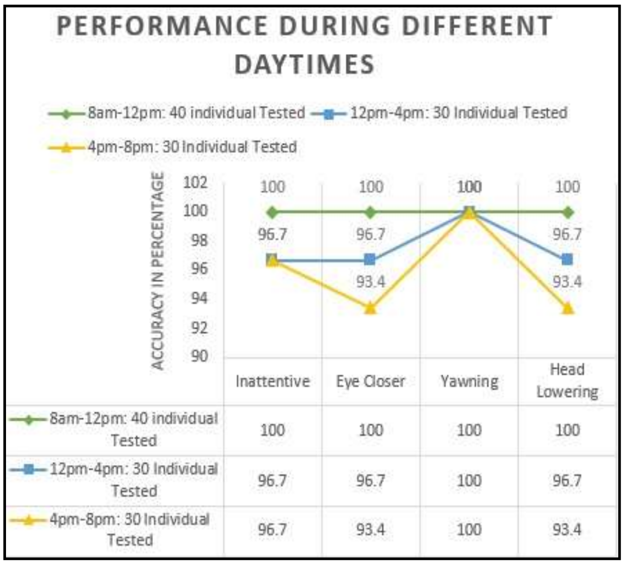
Our proposed system has been implemented in MATLAB version R2017a and run on PC with Intel Core i5 2.40 GHz CPU processor and 8GB RAM. These subjects were asked to simulate typical driving tasks such as looking in front with eyes open for 10 seconds followed by looking in right mirror for 30 ms and left mirror for 45 ms. They were also asked to simulate typical distracted driving behaviour which were looking back to talk to co-passenger for more than 5 seconds, looking outside window for continuous 5 seconds, eye closed for more than 3 seconds. The resolution of these videos were 573 * 814 pixels, and the frame rate was 30 fp Accuracy of detection of all parameters of distraction is calculated and system performance is illustrated on 50 males and 50 females divided in the age groups 18-30 years and 31-above, shown in the Fig.7, in which all five parameter's detection accuracy is shown [13].



	Total Individuals	Inattentive	Eye Closer	Yawning	Head Lowering
Age Group: 18-30	55	100	97.5	100	96.07
Age Group: 31-65	45	95.03	95.03	100	98.33

Figure 7: Performance Measurement.

Also, different times of day were considered for capturing videos to evaluate the detection of all visual cues of inattention and distraction at varied luminance.

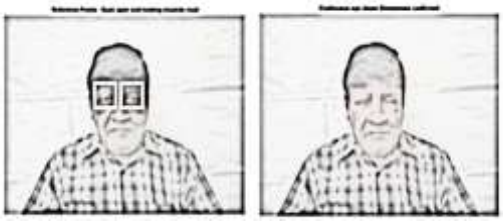


	Inattentive	Eye Closer	Yawning	Head Lowering
8am-12pm: 40 individual Tested	100	100	100	100
12pm-4pm: 30 Individual Tested	96.7	96.7	100	96.7
4pm-8pm: 30 Individual Tested	96.7	93.4	100	93.4

Figure 8: System Performance during Different Daytimes.

It is evident from Fig.8, that in the time slot 4 p.m. – 8 p.m., drowsiness, inattention and head lowering performance got affected. The main reason is decrease in illumination. The driver's face appears to be dark in decreased illumination and bleached out in lightness. But from system's face detection point of view, the visible shade that the system gets to see in the driver's image is outside its programmed accepted range of facial hues. So, system could not detect faces from the images in such cases resulting in false detection of distraction.

Results of eye closure detection algorithm



Reference frame: eyes are open, (b)Drowsiness confirmed Frame: eyes are continuous close for more than 3 seconds.
Figure 9: Eye Closure Detection Results.

Fig.9 shows the results of execution eye closure detection algorithm on drivers. Fig.9(a) is the first image frame called as reference image captured at the start of the system. The valley distances of this image is referred to as reference and used for comparison with next frames as discussed in section 4.4.2. Therefore, based on valley distance, eye state is decided to be open or close. Fig.9(b) is the closed eye frame. The system checks for two conditions, i.e. blinking and drowsiness, as soon as first closed eye frame is encountered. The system has confirmed this frame as drowsy driver case since the consecutive closed eye frames have been encountered continuously for more than 3 seconds. The audio warning alert sound "TIME FOR TEA BREAK" is produced as system confirms drowsiness.

Results of yawning detection algorithm

Figure 10: Results of Yawning Detection.

Fig.10 shows the results of yawning detection algorithm. Here, yawning is detected when distance between the two lips is larger than pre-defined threshold discussed in section 4.4.3.2. Also, by using merge thresholding value, we have distinguished between talking and yawning. Fig.10(a) shows driver's non-drowsy frame. The system checks for such closed mouth in each frame. When, there is encounter of open mouth as shown in Fig.10(b), then algorithm tries to compare it with threshold. If its greater than the threshold the frames detects yawning and raises an alarm sound "STOP YAWNING".

Results of Head Lowering Algorithm

Figure 11: Results of Head Lowering Detection.

Fig.11(a) shows the reference non-drowsy image of driver which is taken as reference. The skin segmentation algorithm has been implemented to calculate the skin pixels in the reference image. Fig.11(b) shows skin segmented head lowered image in which it is evidently visible that the skin pixels are comparatively reduced than Fig.11(a). Hence, this marks the presence of head lowering case and alarm sound is generated "HEAD IS LOWERED, STOP THE CAR", to wake-up the driver.

Results of Inattentiveness detection algorithm

Fig.12 shows the results which are obtained after execution of inattentiveness algorithm on the driver. As discussed in section 3.3, the algorithm decides level of attention based on head pose direction of driver. In Fig.12(a), the driver's head is frontal which implies driver is looking towards road. But, it is evident from Fig.12(b), the driver's is not looking towards road. The direction angle indicates the driver is looking in mirror. The mirror-viewing is not a type of distraction and such non-frontal frame did not last for more than 3 seconds since no driver needs to look in mirror continuously for more than 3 seconds. The system detects the frame as mirror-viewing frame and do not raise any alarm. In Fig.12(c), the driver is talking to co-passenger. The system detects this frame as inattentive. For inattention detection consecutive non-frontal frames are encountered continuously for more than 3 seconds and alarm sound "STAY ALERT" is produced to warn the driver about potential danger due to inattention.

(a) Indicates Reference Frame Captured at Start of System, (b) Driver Is Looking in Mirror, Which Is a Valid Driving Condition, (c) Driver Is Inattentive and "STAY ALERT" Alert Is Raised.

Figure 12: Results of Inattentiveness Detection.

Table 1 shows the accuracy of detection of individual parameters of distraction and the audio alerts generated by the system. In Eq. (2) True positive refers to the individuals whose distractions are correctly classified. False negative refers to the cases in which the individuals distracted are not correctly detected.

$$Accuracy = \frac{TruePositive}{TruePositive + FalseNegative} \quad \ldots (2)$$

Table 1: Accuracy of Distraction Parameter.

Parameter	Accuracy (%)	Audio Alerts
Inattentive	98	Stay Alert
Eye Closer	97	Time for Tea Break
Yawning	100	Stop Yawning
Head Lowering	97	Head Lowered. Stop The Car

Inattentiveness and yawning is efficiently detected by proposed algorithms. Both the algorithms are developed by adapting haar-like features based on Viola Jones proposed method. The algorithm is in-variant to changes in illumination such as under exposure or over-exposure to lighting and these results in greater detection accuracy. The detection of eye closure and head lowering algorithms are also efficient. But in these cases, algorithms are frantically searching for combinations of pixels that could indicate a face. With shadow and underexposure, all the system finds in the image is a patch of darkness. With overexposure, all it finds is a patch of whiteness and glare. The system does not see clearly detect the face in these cases.

6. CONCLUSION

This study presented a vision-based driver monitoring framework for the real-time detection of driver distraction and drowsiness using multiple behavioural cues, including eye closure, yawning, head orientation, and attentiveness assessment. Unlike conventional approaches that rely on a single indicator, the proposed framework integrates complementary visual features to improve the reliability and robustness of driver-state recognition. The use of symmetry-based feature extraction and computationally efficient facial analysis enables real-time operation while reducing system complexity. The experimental evaluation demonstrated high detection accuracy across various distraction and fatigue scenarios, highlighting the effectiveness of the proposed approach in monitoring driver vigilance. From a road-safety perspective, the framework contributes toward the development of intelligent and human-centred Advanced Driver Assistance Systems (ADAS) by providing timely alerts for potentially hazardous driving behaviour. Furthermore, the reliance on interpretable behavioural indicators enhances transparency in the decision-making process, thereby supporting the broader principles of Explainable and Responsible AI in safety-critical transportation applications.

7. FUTURE SCOPE

Future work will focus on incorporating deep learning-based perception models, explainability mechanisms for decision interpretation, bias and fairness evaluation across diverse driver demographics, and multimodal sensor fusion integrating in-cabin cameras, vehicle telemetry, and physiological signals. Such enhancements can further improve the trustworthiness, generalization capability, and real-world deployment readiness of next-generation AI-driven driver monitoring systems.

Acknowledgements: Thank you Astha, Shruti, Samruddhi for helping me in data collection and initial analysis.

REFERENCES

[1]. Manu B. N., “Facial Features Monitoring for Real-Time Drowsiness Detection,” 2016 12th International Conference on Innovations in Information Technology (IIT), pp. 78–81, 2016.

[2]. Yogesh Chellappa, Narendra Nath Joshi, Vaishnavi Bharadwaj, “Driver Fatigue Detection System,” 2016 IEEE International Conference on Signal and Image Processing (ICSIP), pp. 655–660, 2016.

[3]. Meng-Che Chuang, “Estimating Gaze Direction of Vehicle Drivers Using a Smartphone Camera,” The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 165–170, 2014.

[4]. H. Eren, S. Makinist, E. Akin, A. Yilmaz, “Estimating Driver Behavior by a Smartphone,” IEEE Intelligent Vehicles Symposium, Alcalá de Henares, Spain, pp. 234–239, 2012.

[5]. M. Omidyeganeh, A. Javadtalab, S. Shirmohammadi, “Intelligent Driver Drowsiness Detection Through Fusion of Yawning and Eye Closure,” 2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems (VECIMS), 2011.

[6]. Paul Viola, Michael Jones, “Robust Real-Time Face Detection,” Kluwer International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, 2004.

[7]. Dinesh Kadam, Amol R. Madane, Krishnan Kutty, S. V. Bonde, “Rain Streaks Elimination Using Image Processing Algorithms,” Signal & Image Processing: An International Journal (SIPIJ), vol. 10, no. 3, pp. 21–32, Jun. 2019.

[8]. Ashish Kathe, Pranav Teli, Amol R. Madane, “COVID-19 Regulations Check: Social Distancing, People Counting and Mask Wear Check,” 6th International Conference on Data Management, Analytics and Innovation 2022 (ICDMAI 2022), India, pp. 467–484, 14–16 Jan. 2022.

[9]. Priyojit Chakraborty, Amol R. Madane, “Vehicle Queue Management System at Fuel Station,” NIELIT's International Conference on Communication, Electronics and Digital Technologies (NICE-DT'23), New Delhi, India, 10–11 Feb. 2023.

[10]. Amol Raoji Madane, Dr. Ho Chin Kuan, “Convxpf-Net: An Explainable AI Approach for Recognizing Pedestrian Behaviour Using Convolutional XGBoost Pyramid Fractional Network,” International Journal of Advances in Signal and Image Sciences, Nov. 2025.

[11]. Amol Raoji Madane, Dr. Ho Chin Kuan, “Adaptive Multi-Spectral Vision Fusion for Reliable Hazard Identification in Challenging Visibility Conditions,” International Journal of Applied Mathematics, vol. 38, no. 12s, pp. 4241–4257, Dec. 2025. (ISSN: 1311-1728 (printed version), ISSN: 1314-8060 (online version)).

[12]. Amol R. Madane, Dr. Uttam D. Kolekar, “High Frequency Component Based Addition Algorithm of Digital Image Watermarking,” International Journal of Computer Applications, vol. 50, no. 10, pp. 30–34, Jul. 2012. (ISBN: 973-93-80869-59-3) (Impact Factor: 0.821).

[13]. Amol Raoji Madane, Dr. Ho Chin Kuan, “Real-Time High-Speed Autonomous Driving via LiDAR–Event Camera Fusion Using Attention-Enriched Spatio-Temporal Graph Neural Networks,” International Journal of Advances in Signal and Image Sciences, Sep. 2025.