ORIGINAL RESEARCH

A novel algorithmic approach of open eye analysis for drowsiness detection

Nageshwar Nath Pandey[1] · **Naresh Babu Muppalaneni**[1]

Abstract Drowsiness is the crucial reason for road accidents nowadays, as per the available statistics. Several valuable lives may collapse because of this. Such valuable lives can be rescued via the detection of drowsiness at its earlier stage. This paper emphasizes a novel algorithmic approach to recognize the driver's drowsiness at its initial stage with remarkable accuracy via employing computer vision techniques purely. Our proposed work has selected the most noteworthy temporal features of eyes (Eye Aspect Ratio, pupil's center) and head (tip of the nose) to classify the driver's drowsy state more precisely. Further, our developed framework resolved the issue of occluded frames at its pre-processing step via applying the condition of occlusion. Afterward, we have imposed three checks via employing eye aspect ratio, pupil's center, and head (tip of the nose) movement to ensure the correct drowsy state of the driver. Consequently, the performance and accuracy of the overall system have improved in contrast to existing Techniques. Thus, our proposed system has achieved an accuracy of 94.2% in open eye detection.

Keywords Computer vision · Classification · Drowsiness · Deep learning

✉ Nageshwar Nath Pandey
    nageshwar0@gmail.com

    Naresh Babu Muppalaneni
    nareshmuppalaneni@gmail.com

[1] Computer Science and Engineering, National Institute of Technology Silchar, Silchar, Assam 788010, India

1 Introduction

The sign of "drowsy" emphasizes that somebody got shattered or tired at some stage in working hours. Although the term "Drowsy" usually consider as a common in general perspective, in some scenarios, it becomes crucial when somebody is engaged in a job where attention becomes the critical requirement like persons handle the bulky machine in industries or driving a heavy or passenger vehicle, etc. [1, 2]. Drowsiness creates the great trouble for the pedestrians walking on the road or passenger presents in the vehicle. In accordance with the yearly report 2018 over accident analysis in India specified by the Ministry of Road Transport & Highway highlighted that 4,67,044 accidents occurred in states and Union Territories [3]. From statistics of accidental data, this report uncovers the facts that 1,51,417 lives, as well as 4,69,418 people, were injured in these crashes. Of these accidents, 78% have occurred because of the diversion from concentration or drowsy activities of the driver. The yearly comparison between roads crashes showed 0.34% higher concerning the past year, which is also mentioned in this report. Further, the report also showed that from among the total percentage of deaths regarding road accidents worldwide, 11% of deaths were found in India. Further, this report emphasizes that most of the victims were concerned with 18–45 years.

Together with the above discussion, a small number of researches also show that one driver goes in a drowsy state out of 25 drivers at the time of driving [4, 5]. The report also reveals that crashes mostly occurred next to arrival place after a lengthy trip.

According to accidental road data analyzed in the previous paragraph, we can say that driver's act is one of the essential aspects influencing road security. That is why

2200

Int. j. inf. tecnol. (December 2021) 13(6):2199–2208

scrutinizing the driver act related to drowsiness recognition has drawn the attention of researchers towards this field.

Miscellaneous techniques have been evolved for supervising the state of drowsiness in the driver. These techniques can be usually categorized into three classes, i.e., biological feature, vehicle movement indicators, and behavioral feature.

In the biological feature, the symptoms of drowsiness are captured through various signals such as brain signals, heart rate, and nerve impulses. Although biological features-based technique gives the significant result with a higher rate of accuracy but here, it is expected to use extra equipment needed to mount with the body of the driver to acquire these features. This extra equipment creates a hindrance to the everyday driving of the vehicle. Since this technique is invasive and hence, cannot use for the practical purpose [6].

In the second category, i.e., vehicle movement based drowsiness identification involves the vehicle characteristics such as lateral pose, steering wheel movement, bending angle, vehicle pace, pressure on brake or acceleration paddle, lateral displacement, and other signs of vehicle movements [7]. The benefit of this method is that capturing these features is absolutely simple and significant.

Besides these benefits, these techniques have many restrictions like driver expertise, type of vehicle, road shape, and situation. As this technique demands sufficient time to acquire these parameters. Therefore, these methods do not function well in the case of micro-sleep, which means that the driver goes in a sleepy state for a small duration on a linear road without fluctuating signals.

The problems persist in above-discussed methods are addressed by third category of drowsiness detection technique, i.e., behavioral based technique. Here, this technique is non-invasive means no extra equipment is essential to attach with the body of the driver to capture the features as the techniques discussed in the previous paragraph. Consequently, the behavioral-based technique has been famous among researchers because of its lesser cost of organization and ease of use. Primarily computer vision technique is used here to acquire the features via camera only [12–14]. By keeping this view in mind, we have employed this technique in our proposed framework.

As distinct facial cues such as head movement, pupil movement, facial expression, etc., emphasize the different levels of alertness in the driver. Among these cues, eye state identification is a prominent one. Various systems have been developed to recognize the sign of drowsiness based on eye state analysis, i.e., open or close eye [15–17]. These works used typical cameras and did not involve any training data. Our developed system is also configured based on this approach. Here, no training data is required and capable of capturing drowsiness at its early stage in real-time. In this way, it resolved the issue of having a good dataset for training as well as for testing purposes. Our proposed system works on a novel algorithmic approach as per the current demand.

The key objective of our proposed work is to remove the problem of occlusion as not resolve by the available works. Thus improve the accuracy of the proposed model. This proposed work represents the extensive study of drowsiness detection frameworks and develops a highly optimized and reliable model with significant results.

This entire study is divided into five sections: Sect.2 contains a literature review of previous work and its comparison to the proposed work. The proposed technique is detailed in Sect. 3. Further, Sect. 4 includes a result analysis and discussion of the proposed work. Section 5 contains ideas for future work as well as the conclusion of the proposed work.

2 Literature survey

Investigation of existing works has been done under this section. Here, we have incorporated the most influential works from various sources regarding drowsiness detection. Further, the authors examine a variety of publicly accessible datasets concerning national and international domains.

Here, the authors have carried out the investigation of drowsiness detection based on temporal and spatial features that utilize computer vision and deep learning concepts, as shown in Table 1 [8–11].

We have given a novel algorithm for drowsiness detection using open eye state analysis in this present work. Here, our proposed work demands a conventional camera only. Therefore, the model does not require any training data. However, we have considered a small publically available IMM face dataset [18]. In this way, our proposed system is easy to configure, secure and cost-effective.

This is obvious from the survey that every facial cue, such as head rotation, pupil movement, eye blinking and yawning, etc., has its degree of significance in detecting a drowsy state [19]. Among these facial cues, eye state analysis is a prominent one that helps predict the state of alertness better. Several systems have been proposed to identify the level of drowsiness based on the investigation of eye states [20]. From these existing systems, few employed the four-step procedure, i.e., face recognition, localization eye region of interest (ROI), analysis of different states of the eye, and predict level of drowsy state [21]. These existing systems need only an ordinary camera and hence do not involve any training data for experimental purposes.

Table 1 Comparison between previous work and proposed work

Features	Metric	Pros	Cons	Classifier	Accuracy
Temporal feature (proposed work)	Eye Aspect Ratio (EAR), Pupil centre	Execution is very less	Resolve the problem of occlusion	Linear Support Vector Machine (SVM)	94.55%
Temporal feature [8]	Blinking feature of eye	Less training time	- Higher false positive rate reduce the accuracy - Occlusion problem	Linear SVM	65.2%
Spatial + temporal [9]	Eye blinking features	- Training time is less	- Higher false positive rate - Occlusion issue present	Blink duration	93%
Temporal feature [10]	Features of eye, mouth and head-pose	Model tuning is not needed	-Various lighting condition is not used in capturing blinks and all - Occlusion issue present - Used eye features only	Adaboost classifier	88.23%
Temporal feature [11]	Emotional feature	- Under fitting problem solved via Generative Adversarial Network	-Occlusion issue persists - Problem of low illumination	Softmax classifier	87.1%

Few existing works skipped the face identification step [22, 23]. As a result, these systems acquired a higher speed. Subsequently, they had to depend on both training data and particular hardware like illuminators that can be risky from the health point of view due to the direct incidence of light on the eye.

Therefore, we have employed dlib software to recognize the face where hardware like illuminator and all was not needed. Our procedure involves the novel algorithmic approach for eye state analysis for drowsiness detection.

Numerous techniques have been proposed for face identification via NIR [24–26], deep neural network [27, 28], pattern matching [9], Gabor filters [29], skin color [30, 31]. Techniques that support skin color based identification may be limited to a particular ethnic crowd or degrade the performance under different lighting situations.

Further, methods based on the neural network [27] for face detection give better results but fail in terms of more processing time due to network training, which does not suit the requirement of early drowsiness detection.

Afterward, the existing framework [9] employed the concept iris-sclera pattern-based approach to detect open eye drowsiness detection performs better. However, it fails when the eye has already been suffering from defects like squinting or when the driver sees in the left or right direction extremely. In this scenario, the symmetry concept of sclera and iris will not work better to determine the state of the open eye, and hence the system will start mispredicting the cases of drowsiness. Further, this existing system also could not resolve the issue of the occluded frame as well as identification of eye region of interest under dim lighting conditions.

Although, the model that has employed the techniques of Gabor filter [29] gives the better result in the sense of accuracy but fails in terms of more time to train the model as well requirement of a suitable and significant amount of dataset for training and testing purpose which further a challenging task.

Likewise, the existing frameworks [32] that implement the technique of Viola-Jones based on Haar cascade feature gives the superior result with a faster rate of face identification but require dataset for training as well as testing purpose. Such a system is also does not capable of the suite the problem of occlusion.

In a comparison of previous works, our developed model has contributed in the following ways:

- In our developed work, we have employed the most prominent features, i.e., pupil movement and eye aspect ratio, which reduces the processing time and enhances the accuracy.
- In our proposed work, we have removed the issue of occluded frames via using the criterion of occlusion at its pre-processing step and further increase the accuracy.
- Our developed system resolved the challenge of having a good dataset for the training and testing purpose of the model and hence reduced the processing time of the overall model.
- In our proposed framework, we have created a reliable system that will not only identify drowsy states but could also capture the moment when the driver is departed from routine driving.

3 Proposed methodology

Since this is obvious from the techniques examined in Table 1, no one has given a proper solution to suit the challenges of occlusion that arise due to head rotation. In our developed framework, we have resolved the issue of occlusion by removing the defected frames via satisfying the criterion of separation of the pupil's center and horizontal length of an eye using the computer vision technique. In this way, our proposed framework upgrades the accuracy of the overall system. The outline of our developed framework is exposed in Fig. 1.

3.1 Dataset

As per the rigorous literature survey, we have found that most of the dataset was not wholly suitable for our experimental purpose. However, we have chosen the rarest IMM face dataset, which was well-annotated and satisfied our need for experimentation, i.e., detection of drowsiness with an open eye, to a certain extent because it contains all the images and video with an open eye. The difficulty with this dataset was that it did not persist any image or video where the subject was wearing the glasses. Therefore, we have included a few such images and videos from other online sources in order to maintain the versatility of this dataset.

In order to validate our proposed system, we have carried out a couple of experiments. In the first phase, we

Fig. 1 Flow chart of proposed work

experimented on only color images present in the IMM face dataset [18]. This dataset contains 222 color images along with 18 grayscale images. These images have been contributed by 37 different subjects, with 6 images belonging to each subject. These six images related to each subject are labeled based on facial expressions ("normal," happy"), lighting circumstances as well as facial rotation, i.e., "30 degrees movement to the subject's right," "30 degrees movement to the person's left" and so on. These captured images are with resolution 640×480 without bearing glasses. In the next phase of the experiment, we have considered the six video sequences of a different subject with the various angle of head movement. These videos have been captured under various lighting conditions through the ordinary camera with a resolution of 640×480 at the rate of 30 frames/sec.

To build the multipurpose dataset, we have included the persons related to the diverse civilization, colors, and different environments. Further, the images, as well as videos, were acquired at a variety of angles under the different scenarios.

3.2 Facial feature extraction via Dlib

From the rigorous survey of the literature, we have found that several techniques are available for extracting the facial cues from the face. Iqbal et al. [33] have designed a framework based on a deep learning technique called Visual Geometry Group network (VGGFace) for recognizing the face in the image. This framework was having an analogous arrangement like a convolution neural network (CNN). This model eliminates the data in a hierarchical fashion. This model captures the face with a higher accuracy rate, i.e., 99.77%, but took more time to identify a face in the image. Thus, this model could not satisfy the criterion of early drowsiness detection. Likewise, Kambi et al. [34] have proposed a model that resolves the issue, such as huge alteration in facial expression as well as pose. This method contains the dual approach of LBP and k-Nearest Neighbor (k-NN) techniques. Further, this technique was not capable of capturing the faces in low illumination conditions. In order to resolve these problems, we have employed the concept of key-points-based techniques of face identification where instead of recognizing the face ultimately, the only region of interest in the face is captured via inbuilt landmark points [35]. Thus it reduces the overall processing time and hence satisfies the demand for early drowsiness detection with remarkable accuracy.

Thus in our proposed work, we have employed the Dlib facial feature predictor package available in OpenCV libraries for the identification of the region of interest in the face [36]. Here, the Dlib package consists of a Histogram of Oriented Gradient (HOG) feature descriptor for
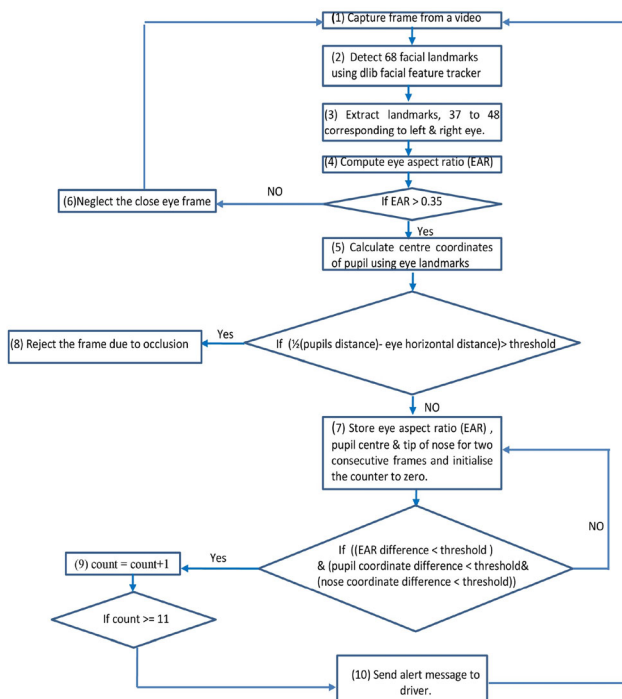
detecting region of interest in the face as well as Linear Support Vector Machine (SVM) as a classifier [37]. This Dlib package is associated with 68 facial landmarks which makes easier to extract the numerous region of interest (ROI) from the face. According to our requisite, we have obtained the landmarks sequencing from 37 to 48 to trace the pair of eyes to compute the center of pupil and eye aspect ratio (EAR) for imposing condition of drowsiness in open eye situation as shown in Fig. 2. These landmarks are extracted from each frame. Afterward, we obtained the coordinates, i.e., *x* and *y*, of these extracted landmarks to compute the value of EAR and center of the pupil for further processing.

3.3 Pre-processing technique

In our experimentation process, we have analyzed that it was challenging to capture the region of interest, such as eyes, mouth, and everything, in the presence of poor illumination, especially at night. As a result, we used the Histogram equalization approach, which evenly distributes intensity values throughout the frame [38, 39]. Thus, the dispersion of unequal light in each frame is reduced. Thereafter, we utilized the Gamma correction approach to increase contrast across the frame by performing a non-linear adjustment between the input and output mapped values [40].

3.4 Eye aspect ratio (EAR)

The term eye aspect ratio can be characterized as the ratio of perpendicular distance to the flat distance of an eye. In pursuance of the Dlib software package, eyes are designated with six inbuilt facial landmarks indexed from p1 to p6, as shown in Fig. 3.
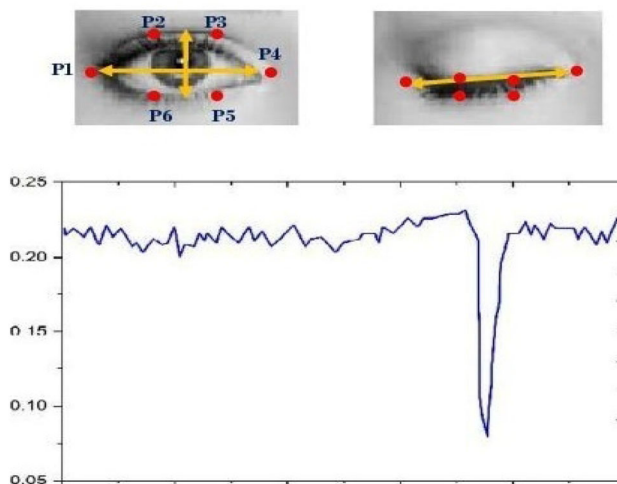

Fig. 3 Eye aspect ratio

Here, these landmarks participated in the calculation of the eye aspect ratio. In the computation of eye aspect ratio, initially, we determine the perpendicular distance via the evaluation of the midpoints through the coordinates of a couple of landmarks (p2, p3) and (p5, p6). Further, flat distance is computed from the coordinates of a pair of landmarks (p1, p4) using the Euclidian distance formula. Thus, the value of EAR is computed using Eq. (1) as follows:

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \tag{1}$$

3.5 Centre of pupil

As in pre-processing step of our framework, occluded frames are neglected based on certain conditions. Further, to establish this condition, the center of pupils is one of the

Fig. 2 ROI selection using 68 facial landmarks
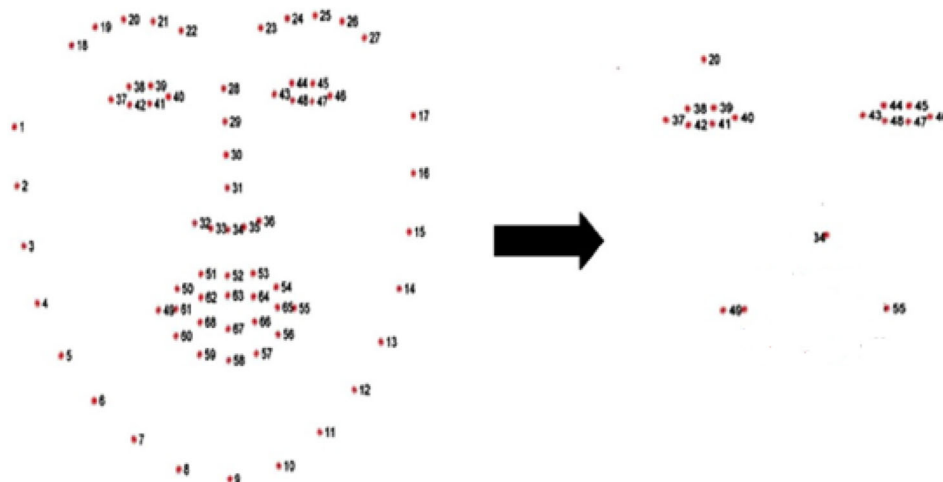
essential parameters to be determined. Therefore, firstly we will compute the center of pupils.

In the initial stage, we have cropped eye region of interest (ROI) via extracting coordinates of relevant facial landmarks for localizing the center of pupils from each frame of a video. Further, gradient-based concept, i.e., pupil region is darker than the neighbouring region, has been employed to localize the center of the pupil. Afterward, we used a Gaussian blur filter with size 7×7 to reduce noise from each cropped eye ROI of a frame. After that, we have localized all the contours and pick the largest one based on the area of contour. Thus, the obtained largest contour will be the pupil. Further, we have enclosed the pupil through a rectangular bounding box. After that, we have computed the center of a pupil from the width and height of the bounding box as given in Eqs. (2) and (3) and from Fig. 4 as follows:

$$x_p = x + w/2 \tag{2}$$

$$y_p = y + h/2 \tag{3}$$

where, w, width of the bounding box containing pupil' h, height of the bounding box containing pupil.

Here, we have used boudingRect () function to recognize (x, y, w, h) coordinates.

3.6 Condition of occlusion

The proposed framework of Yingyu et al. [36] emphasizes that a frame is said to be occluded when half of the line connecting the center of the pupil becomes equal to the horizontal length of an eye, i.e., either left or right eye. Consequently, we have imposed this condition to identify whether a frame is occluded or not at the initial stage of processing. Since, we have localized the center of pupil in the earlier Sect. 3.3, Now, we are capable of computing the length of the line connecting the pupil's center from the coordinates of the center's of the pupil's by utilizing simple distance formula as presented in Eqs. (4) and (5) as well as from the Fig. 5 as follows:
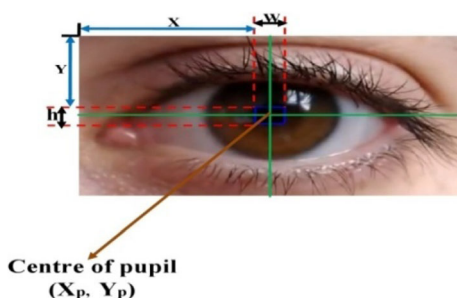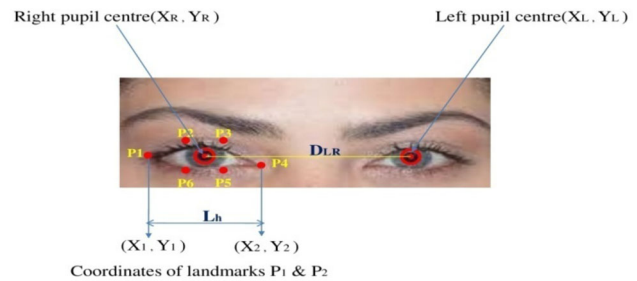

Fig. 5 Computation of condition of occluded frame

$$D_{LR} = \sqrt{(x_L - x_R)^2 - (y_L - y_R)^2} \tag{4}$$

$$L_h = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2} \tag{5}$$

where (x_L, y_L) and (x_R, y_R) are the pupil center's for left

and right eye. Further, $(x_1, y_1) \& (x_2, y_2)$ are the coordinates of landmarks related to right eye, D_{LR} is the length of line connecting pupil centre's of left and right eye, L_h is the horizontal length of an eye.

Thereafter, we have acquired the horizontal distance of an eye via extracting the coordinates of landmarks 37 and 40 concerning to right eye or 43 and 46 related to left eye as shown in Fig. 5. Afterward, we have checked the condition of occlusion as per the Algorithm 1 as follows:

Algorithm 1: *Compute Defected Frame*(L_h, D_{LR}, *threshold*)

Input: L_h, D_{LR} and threshold value from a frame
Output: Corrected frame

1: **If** $(L_h - \frac{1}{2} D_{LR})$ < threshold **then**
2:    Accept the frame for feature extraction
3: **Else:**
4:    Reject the frame due to defect
5: **End if**

Fig. 4 Centre of pupil detection

Further, the complete processing of this proposed model is summarized via the novel algorithmic approach as given in Algorithm 2 as follows:

Algorithm 2: *Compute Drowsy State (Frame, thresh₁, thresh2, thresh 3)*

Input: *video frames and thresh₁ and thresh₂, thresh3*
/* Here, thresh₁ is the threshold value related to eye aspect ratio, thresh2 is the threshold value related to pupil's center, thresh3 is the threshold value related to head movement */
Output: *Drowsy or Alert*

1: Capture the i*th* frame of a video
2: Call *Compute Defected Frame*$(L_h, D_{LR}, threshold)$
3: If frame is not defected **then**
4:      for i = 1 to n
5:           Calculate E = $EAR_{i+1} - EAR_{i+1}$
6:           Calculate $D_1 = |(x_{i+1}, y_{i+1}) - (x_i, y_i)|$
7:           Calculate $D_2 = |(x_{i+1}, y_{i+1}) - (x_i, y_i)|$
/* Here, E is the difference between eye aspect ratio of two consecutive frame, D_1 is difference between the pupil centers in two consecutive frame i.e. i*th* & i+1*th* frame, D_2 is the difference between nose tip in two consecutive frame.*/
8: Initialize counter ctr =0
// ctr counts the pair of consecutive frame
9:           if (E< *thresh₁ and* D_1 *< thresh₂ and* D_2 *= thresh 3*)
10:                ctr = ctr +1;
11:                if (ctr>= 11) **then**
12:                     send alert message to the driver
13:                **Else:**
14:                     process the next incoming frame

---

4 Result and discussion

In order to highlight the outcome for categorization of alert as well as drowsy states of driver based on the open eye in our proposed model, we have employed the concept of counting the number of frames for a particular duration. Stepwise processing of this complete concept is shown in Fig. 6 as follows.

Further, we have shown the result of our model in the terms of Accuracy, Precision and Recall based on parameters of true positive (TP), true negative (TN), False positive (FP) and False negative (FN) which is obtained from the confusion metrics shown in Fig. 9. Here, the formulae used for the computation of Accuracy, Precision and Recall from the Eqs. (6), (7) and (8) as follows:
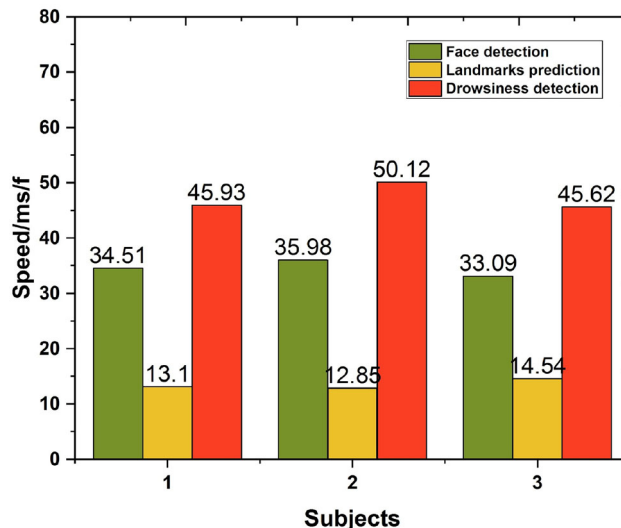
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (6)$$

Fig. 6 Speed of algorithm

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (8)$$

Figure 7 demonstrates the computing speed of individual algorithm during the analysis of videos of three subjects. The average speed of each algorithm can be simply estimated using this graph. As indicated in Table 2, we compared the performance of prior work to our proposed work in terms of accuracy, precision, and recall. Our suggested approach achieves superior outcomes when compared to existing models, particularly those that are linked to temporal features, as this temporal characteristic is also used in our model.

The outcomes of our developed model demonstrate the robust categorization of two classes of drowsiness, i.e., alert and drowsy, with improved accuracy, i.e., 94.2% in comparison to the available model [9], which shows the accuracy of 93%. As from the rigorous review of model [9], we have drawn the conclusion that the symmetry concept utilized in this model was not capable of determining the open state of eyes accurately in the situation when eye has already been suffering from defect like squinting or driver is seeing either left or right directions extremely for a long time. Thus, this system shows the number of false cases of drowsiness prediction, and hence, degrades the performance of overall system for the existing work [9]. To resolve the issues for such cases of false prediction, we have imposed the condition of occlusion on each incoming frames and neglected such a frame at its pre-processing step of our model. In this way, we have achieved the accuracy of 94.2% for our temporal feature-based model compared to existing work where total

2206

Int. j. inf. tecnol. (December 2021) 13(6):2199–2208
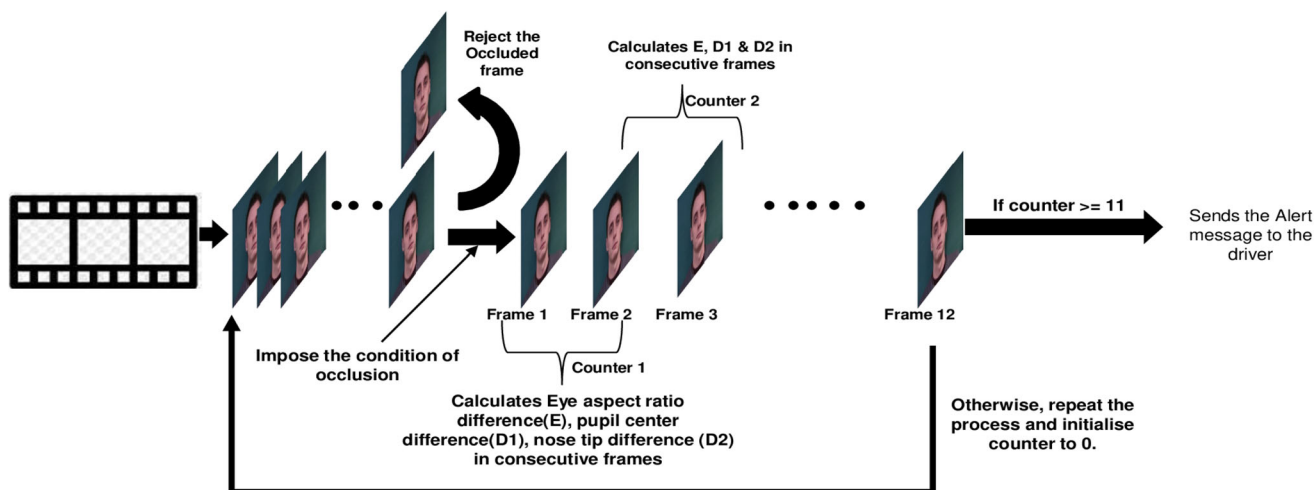
Fig. 7 The complete processing of proposed work

Table 2 Performance analysis of few drowsiness detection techniques

Previous work	Accuracy (%)	Precision	Recall
Temporal feature [8]	65.2	65.35	63.28
Spatial + temporal [9]	93	93.38	91.52
Spatial feature [10]	88.2	86.5	89
Temporal feature [11]	87.1	88.6	87.50
Proposed work (temporal feature)	94.55	95	94.05

accuracy acquired was 93% [9]. On the other hand, the existing work [9] could not resolve the problem of drowsiness detection correctly for the frames captured under dim lighting conditions, especially during the night hours. In comparison, our developed framework has implemented the concept of Histogram equalization and Gamma correction techniques to resolve the issue of a frame captured under dim lighting conditions, especially during night hours and further enhanced the performance of overall system. Instance images of experimental result are shown in Fig. 8, and result statistics in the form of confusion metrics are shown in Fig. 9.

In our proposed work, we have purely employed the computer vision technique, which captured temporal features of an eye (Eye aspect ratio, center of pupil) and mouth (Tip of the nose) for performing three levels of verification to identify the state of drowsiness correctly. Further, our developed system does not require any training data like the existing work [8, 11], and hence the processing time of our developed framework is very less, i.e., 31.74 ms/frame approximately in comparison of existing works [8, 11] where several hours are elapsed in training as well as in testing time. This advantageous point enhanced the robustness of our model because time plays a crucial role in drowsiness detection because drowsiness detection at its earlier stage gives enough time to the driver to

recover from the drowsy state to rescue herself/himself from the accident that is going to happen.

5 Conclusion and future work

The practices utilized in our developed framework are genuinely based on temporal features of an eye (Eye Aspect Ratio, pupil movement) and head (Tip of the nose) movements. Further, we have imposed three levels of checks via using the above-mentioned features through a novice algorithmic approach that shows the robustness of our developed framework. In comparing existing work, we have utilized another novel algorithmic approach to diminish the issue caused by the occluded frame at the pre-processing stage of the experimentation process. A pure computer vision-based technique has been employed to upgrade the accuracy and performance and reduce the processing time. The model does not require any training data, and hence time elapse in training is saved. Thus, the outcomes achieved from the practices reveal that our developed model is better in comparison to existing models.

In support of future work, one can try some additional significant features like the posture of hands, legs, and all to build a more robust model compared to existing models.
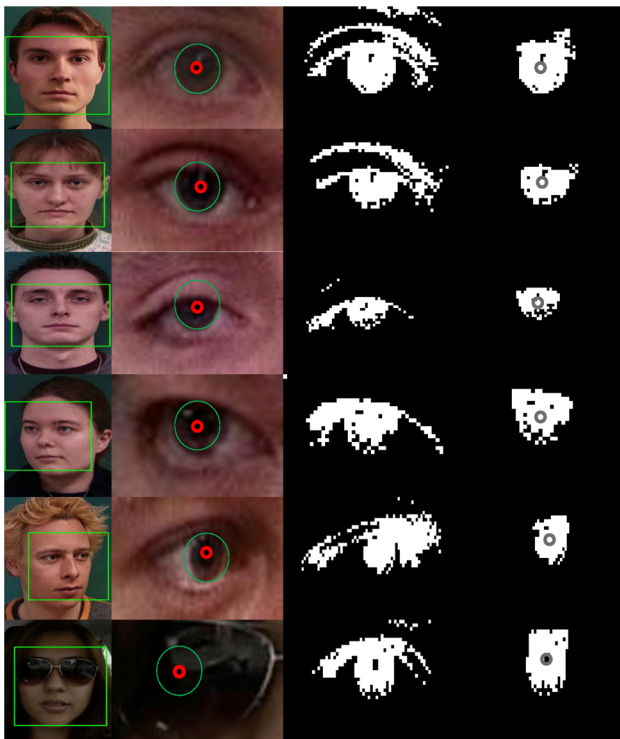
Fig. 8 Instance images from IMM database and online source. First column is images with face detection. Second column contains images with pupil center and iris circle. Third column is inverse binary images of detected eyes. Fourth column contains images after applying Gaussian Filter over inverse binary image

Fig. 9 Confusion matrices for model

At some stage in the experimentation process, we have realized the difficulty of localizing the pupil's center due to the extreme posture of the head in either direction. Consequently, one can employ neural network-based techniques like MT-CNN, Face-Net, and all to capture a face and other regions of interest instead of dlib based technique. Although these techniques will capture the face and other ROI with greater accuracy, they raise the problem of more time elapsed to capture these features. Therefore, it will be essential to compensate time with accuracy with some optimized approach.

Further, we realized the problem in setting the value of threshold to impose a condition of occlusion due to dissimilarity in the dimension of eyes of each person. As a result, there is a necessity to normalize the data in this scenario to upgrade the whole performance of the model. In addition to this, we have also observed that our system was not capable of recognizing the driver's eye correctly under sunglasses as well as sudden changes in light as may arise during vehicle on the road. So, one can try to resolve this issue in future work with help of some robust technique. The dataset we used for our study was rarest and satisfactory but not up to the mark because it did not use cases of subjects bearing glasses and all. That is why; we have considered the video and image data from other sources such as online mode. Thus, one can create his/her dataset that should be versatile and relevant for this study.

Declarations

Conflict of interest The authors declare that there is no conflict of interest in this paper.

References

1. Sadeghniiat-Haghighi K, Yazdi Z (2015) Fatigue management in the workplace. Ind Psychiatry J 24(1):12. https://doi.org/10.4103/0972-6748.160915
2. Jiang L, Xie W, Zhang D, Gu T (2021) Smart diagnosis: deep learning boosted driver inattention detection and abnormal driving prediction. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2021.3103852
3. Road Accidents in India 2018. https://morth.nic.in/sites/default/filesAccidednt.pdf, pp. 1–125, Accessed 2 March 2021
4. Wheaton AG, Shults RA, Chapman DP, Ford ES, Croft JB (2014) Drowsy driving and risk behaviors—10 states and Puerto Rico, 2011–2012. MMWR Morb Mortal Week Rep 63(26):557
5. Wheaton AG, Shults RA, Chapman DP, Ford ES, Croft JB (2013) Drowsy driving 19 states and the district of Columbia, 2009-2010. MMWR Morb Mortal Wkly Rep 63:1033
6. Wei CS, Wang YT, Lin CT, Jung TP (2018) Toward drowsiness detection using non-hair-bearing EEG-based braincomputer interfaces. IEEE Trans Neural Syst Rehabil Eng 26(2):400–406. https://doi.org/10.1109/TNSRE.2018.2790359
7. Cui Y, Xu Y, Wu D (2019) EEG-based driver drowsiness estimation using feature weighted episodic training. IEEE Trans Neural Syst Rehabil Eng 27(11):2263–2273. https://doi.org/10.1109/TNSRE.2019.2945794
8. Ghoddoosian R, Galib M, Athitsos V (2019) A realistic dataset and baseline temporal model for early drowsiness detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. http://www.cv-foundation.org/
9. Panicker AD, Nair MS (2017) Open-eye detection using iris–sclera pattern analysis for driver drowsiness detection. Sādhanā 42(11):1835–1849. https://doi.org/10.1007/s12046-017-0728-3
10. Akrout B, Mahdi W (2021) A novel approach for driver fatigue detection based on visual characteristics analysis. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-021-03311-9
11. Hu Y, Lu M, Xie C, Lu X (2019) Driver drowsiness recognition via 3d conditional gan and two-level attention bi-lstm. IEEE Trans Circuits Syst Video Technol 30(12):4755–4768. https://doi.org/10.1109/TCSVT.2019.2958188
12. Maior CBS, das Chagas Moura MJ, Santana JMM, Lins ID (2020) Real-time classification for autonomous drowsiness

2208

Int. j. inf. tecnol. (December 2021) 13(6):2199–2208

detection using eye aspect ratio. Expert Syst Appl 158:113505. https://doi.org/10.1016/j.eswa.2020.113505

13. Khan MQ, Lee S (2019) A comprehensive survey of driving monitoring and assistance systems. Sensors 19(11):2574. https://doi.org/10.3390/s19112574

14. Verma KK, Singh BM, Dixit A (2019) A review of supervised and unsupervised machine learning techniques for suspicious behavior recognition in intelligent surveillance system. Int J Inf Technol. https://doi.org/10.1007/s41870-019-00364-0

15. Pandey NN, Muppalaneni NB (2021) Real-time drowsiness identification based on eye state analysis. In 2021 international conference on artificial intelligence and smart systems (ICAIS). IEEE, pp 1182–1187. https://doi.org/10.1109/ICAIS50930.2021.9395975

16. Raju VSS, Belwal M (2021) Driver drowsiness detection. Computer networks and inventive communication technologies. Springer, Singapore, pp 975–983

17. Selvakumar K, Jerome J, Rajamani K, Shankar N (2016) Real-time vision based driver drowsiness detection using partial least squares analysis. J Signal Process Syst 85(2):263–274. https://doi.org/10.1007/s11265-015-1075-4

18. IMM database. http://www2.compute.dtu.dk/~aam/

19. Ramzan M, Khan HU, Awan SM, Ismail A, Ilyas M, Mahmood A (2019) A survey on state-of-the-art drowsiness detection techniques. IEEE Access 7:61904–61919. https://doi.org/10.1109/ACCESS.2019.2914373

20. Savaş BK, Becerikli Y (2020) Real time driver fatigue detection system based on multi-task ConNN. IEEE Access 8:12491–12498. https://doi.org/10.1109/ACCESS.2020.2963960

21. El Kaddouhi S, Saaidi A, Abarkan M (2017) Eye detection based on the Viola-Jones method and corners points. Multimedia Tools Appl 76(21):23077–23097. https://doi.org/10.1007/s11042-017-4415-5

22. Gong W, Tan Y, Tai Y (2020) Hierarchical HMMs on eyes for driver drowsiness detection. In: 2020 Chinese automation congress (CAC). IEEE, pp 3328–3333. https://doi.org/10.1109/CAC51589.2020.9327367

23. Cho SW, Baek NR, Kim MC, Koo JH, Kim JH, Park KR (2018) Face detection in nighttime images using visible-light camera sensors with two-step faster region-based convolutional neural network. Sensors 18(9):2995. https://doi.org/10.3390/s18092995

24. Pandey P, Tyagi AK, Ambekar S, Ap P (2020) Skin segmentation from nir images using unsupervised domain adaptation through generative latent search. arXiv preprint

25. Li SZ, Chu R, Liao S, Zhang L (2007) Illumination invariant face recognition using near-infrared images. IEEE Trans Pattern Anal Mach Intell 29(4):627–639. https://doi.org/10.1109/TPAMI.2007.1014

26. Nguyen DT, Pham TD, Lee YW, Park KR (2018) Deep learning-based enhanced presentation attack detection for iris recognition by combining features from local and global regions based on NIR camera sensor. Sensors 18(8):2601. https://doi.org/10.3390/s18082601

27. Mehendale N (2020) Facial emotion recognition using convolutional neural networks (FERC). SN Appl Sci 2(3):1–8. https://doi.org/10.1007/s42452-020-2234-1

28. Balayesu N, Kalluri HK (2020) An extensive survey on traditional and deep learning-based face sketch synthesis models. Int J Inf Technol 12(3):995–1004. https://doi.org/10.1007/s41870-019-00386-8

29. Rani PI, Muneeswaran K (2017) Recognize the facial emotion in video sequences using eye and mouth temporal Gabor features. Multimedia Tools Appl 76(7):10017–10040. https://doi.org/10.1007/s11042-016-3592-y

30. Mutneja V, Singh S (2019) Modified Viola-Jones algorithm with GPU accelerated training and parallelized skin color filtering-based face detection. J Real-Time Image Proc 16(5):1573–1593. https://doi.org/10.1007/s11554-017-0667-6

31. Mutneja V, Singh S (2018) GPU accelerated face detection from low resolution surveillance videos using motion and skin color segmentation. Optik 157:1155–1165. https://doi.org/10.1016/j.ijleo.2017.11.188

32. Murphy TM, Broussard R, Schultz R, Rakvic R, Ngo H (2017) Face detection with a Viola-Jones based hybrid network. IET Biometr 6(3):200–210. https://doi.org/10.1049/iet-bmt.2016.0037

33. Iqbal M, Sameem MSI, Naqvi N, Kanwal S, Ye Z (2019) A deep learning approach for face recognition based on angularly discriminative features. Pattern Recogn Lett 128:414–419. https://doi.org/10.1016/j.patrec.2019.10.002

34. Kambi Beli IL, Guo C (2017) Enhancing face identification using local binary patterns and k-nearest neighbors. J Imaging 3(3):37. https://doi.org/10.3390/jimaging3030037

35. Kortli Y, Jridi M, Al Falou A, Atri M (2020) Face recognition systems: a survey. Sensors 20(2):342. https://doi.org/10.3390/s20020342

36. Kazemi V, Sullivan J (2014) One millisecond face alignment with an ensemble of regression trees. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1867–1874. https://openaccess.thecvf.com/content_cvpr_2014/html/Kazemi_One_Millisecond_Face_2014_CVPR_paper.html

37. Chandra MA, Bedi SS (2018) Survey on SVM and their application in image classification. Int J Inf Technol. https://doi.org/10.1007/s41870-017-0080-1

38. Rao BS (2020) Dynamic histogram equalization for contrast enhancement for digital images. Appl Soft Comput 89:106114. https://doi.org/10.1016/j.asoc.2020.106114

39. Gangonda SS, Patavardhan PP, Karande KJ (2021) VGHN: variations aware geometric moments and histogram features normalization for robust uncontrolled face recognition. Int J Inf Technol. https://doi.org/10.1007/s41870-021-00703-0

40. Ju M, Ding C, Zhang D, Guo YJ (2018) Gamma-correction-based visibility restoration for single hazy images. IEEE Signal Process Lett 25(7):1084–1088. https://doi.org/10.1109/LSP.2018.2839580