Driver Drowsiness and Alert System using Image Processing & IoT

Ms. Swathi R[1]

[1]*Researcher, Bengaluru, Karnataka, India*

---***---

Abstract - *The majority of information is relayed by the eyes, which are a paramount part of the body. When an driver or person is weary, countenances such as blinking, yawning rate, and face tilt change from those in a mundane state. We propose a Driver-Drowsiness Detection System in this effort, which employs video clips to monitor drivers' lassitude status, such as yawning, eye closure length, and head tilt position, without having them to carry sensors on their bodies. We are utilizing a face-tracking algorithm to amend tracking reliability due to the constraints of precedent methodologies. To distinguish facial areas, we utilized a technique predicated on 68 key points. Then we assess the passengers' health utilizing these areas of the head. By integrating the eye, mouth, and head, the Driver-Drowsiness Detection Method can utilize a fatigue alarm to alert the driver.*

Key Words: Image Processing, Drowsiness, eye movements, facial movements

1. INTRODUCTION

Increased demand for modern mobility has necessitated faster car-park growth in recent years. The automobile is now a necessary form of mobility for most people. In 2017, 97 million vehicles were sold worldwide, up 0.3 percent from 2016. In 2018, the total number of automobiles in use worldwide was estimated to reach over 1 billion. Although the vehicle has altered people's lifestyles and made daily tasks more convenient, it is also linked to several negative consequences, such as road accidents. According to the National Highway Traffic Safety Administration, there were 7,277,000 traffic accidents in 2016, with 37,461 people killed and 3,144,000 people injured. Fatigued driving was responsible for roughly 20% to 30% of the traffic accidents in this study. As a result, fatigued driving is a major and hidden risk in traffic collisions. The fatigue-driving-detection technology has become a prominent research area in recent years.

Positivist and interpretivist detection methods are the two types of detection procedures. A driver must engage in the subjective detection method's evaluation, which is linked to the driver's subjective impressions through procedures such as self-questioning, evaluation, and filling out questionnaires. These statistics are then utilised to estimate the number of cars driven by weary drivers, allowing them to better arrange their timetables. The objective detection approach, on the other hand, does not require drivers' feedback because it analyses the driver's physiological condition and driving-behaviour parameters in real time. The information gathered is utilised to determine the driver's level of weariness. In addition, objective detection is divided into two categories: contact and non-contact. Non-contact is less expensive and more convenient than contact because the system does not need Vision - based technology or a sophisticated camera, allowing the gadget to be used in more cars.

The non-contact approach has been widely employed for fatigue-driving detection due to its ease of installation and low cost. Concentration Technology and Smart Eye, for example, use the motion of the driver's eyes and the position of the driver's head to estimate their fatigue level. Alerting concern person also is one of the techniques make system more reliable in real world.

1.1 EXISTING SYSTEM

Changes in the eye-steering correlation can indicate distraction, according to the current system. Eye movements linked with road scanning procedures have a low eye steering connection, as shown by autocorrelation and cross-correlation of horizontal eye position and steering wheel angle. On a straight road, the eye steering correlation will regulate the connection. Because of the straight route, the steering motion and eye glances had a low association. This system's goal is to identify driver distraction based on the visual behaviour or performance; therefore, it's used to describe the relationship between visual behaviour and vehicle control for that purpose. This approach assesses the eye-steering correlation on a straight road, assuming that it will have a qualitatively and quantitatively different connection than a curving road and that it will be susceptible to distraction. On curving roads, a high eye steering connection linked with this process has been discovered in the visual behaviour and vehicle control relationship, which reveals a basic perception-control mechanism that plays a major role in driving.

In computer vision, object recognition tracking is a critical issue. Human-computer interaction, behavior recognition, robotics, and monitoring are just a few of the disciplines where it can be used. Given the initial state of the target in the previous frame, object recognition tracking predicts the target position in each frame of the image sequence. The pixel connection between adjacent frames of the video sequence and movement changes of the pixels, according to Lucas, can be used to monitor a moving target. This technique, on the other hand, can only recognize a medium-sized target that moves between two frames. Bolme presented the Minimal Outcome Sum of Squared Error (MOSSE) filter, that can provide robust correlation filters to follow the object, based on recent improvements in the correlations filter in computer vision. Although the MOSSE has a great computing efficiency, its algorithm precision is limited, and it can only analyze one channel's grey information [1].

Tracking of visual objects. In computer vision, visual object tracking is a critical issue. Human-computer interaction, behaviour recognition, robotics, and surveillance are just a few of the disciplines where it can be used. Given the initial state of the target in the previous frame, vision - based monitoring estimates the target position in each frame of the image sequence. The pixel relationship between subsequent frames of the video sequence and displacement changes of the pixels, according to Lucas, can be used to monitor a moving target. This technique, on the other hand, can only recognise a medium-sized target that moves between two frames [3].

2. OBJECTIVE

The driver's eye is tracked using a real-time eye-tracking system. When a motorist is tired or preoccupied, he or she has a slower reaction time in different driving circumstances. As a result, the chances of an accident will increase. There really are three ways to identify tiredness in a driver. The first is physiological changes in the body, such as pulse rate, brain messages, and cardiac activity, which a wearable wristband system can detect. The suggested eye tracking system achieves the second method through behavioural metrics such as unexpected nodding, eye movements, yawning, and blinking.

The developed system's goal is to achieve the five primary points:

a. Affordable: The technologies must be affordable, as cost is one of the most important considerations during the design phase.

b. Portable: The systems to be portable and easy to install in different vehicles models.

c. Secure: The system's security is ensured by placing each component in the proper area.

d. Quick: Because an accident occurs in a matter of seconds, the reaction and process time to react in the event of a driver's urgency is among the most important variables.

e. Accurate: Because the system must be precise, the most precise algorithms have indeed been chosen.

3. METHODOLOGY

One of the major reasons of road accidents in real world has solution now. The system is one step towards safeguarding precious lives by avoiding accidents in real world. Proposed system is based on DLIB & SOLVE PNP Models.

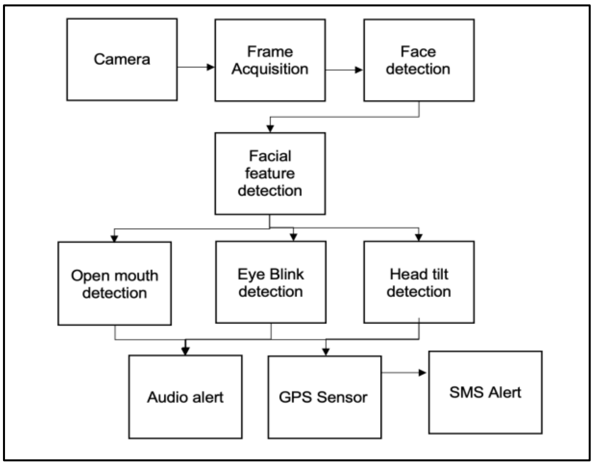

Fig- 1 : Block Diagram of Proposed System

Camera is initialized at the beginning; frame acquisition is done in the next step. After Frame is acquired, face is detected in the frame. In the next step we detect facial features, position of eye, mouth & head is identified.

Following is the formulae to calculate Eye Aspect Ratio:

$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Following is the formulae to calculate Mouth Aspect Ratio:

$$\text{MAR} = \frac{\|p_2 - p_8\| + \|p_3 - p_7\| + \|p_4 - p_6\|}{2\,\|p_1 - p_5\|}$$

We calculate the head Position using PnP i.e. Perspective-end-Point by detecting 3D facial points:

a. Tip of the nose: (0.0, 0.0, 0.0)
b. Chin: (0.0, -330.0, -65.0)
c. Left corner of the left eye: (-225.0f, 170.0f, -135.0)
d. Right corner of the right eye: (225.0, 170.0, -135.0)
e. Left corner of the mouth: (-150.0, -150.0, -125.0)
f. Right corner of the mouth: (150.0, -150.0, -125.0)

After calculating EAR, MAR & PnP, system alerts through speaker if it detects eyes closed for more than 5 secs and yawning so much. System also alerts driver when his head is not straight. As additional security measure system also sends SMS alert to concern person including GPS location fetched via GPS sensor.

4. IMPLEMENTATION

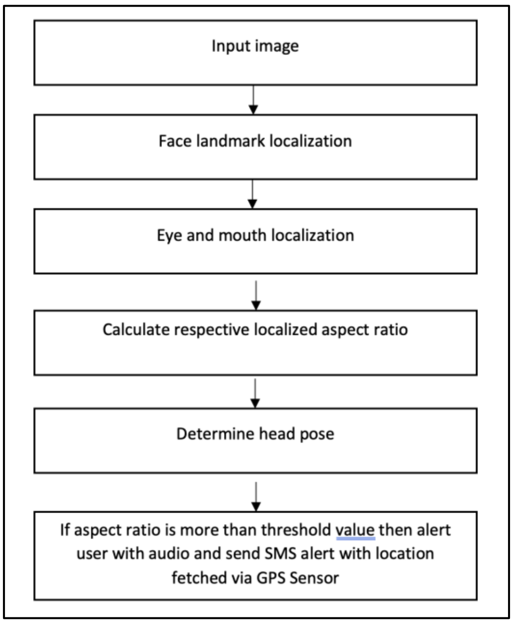

Fig- 2 : Implementation of Proposed System

Facial landmarks are being used to identify and represent important facial features, such as:

- Eye
- Eyebrows
- Nose
- Mouth
- Jawline

Face alignment, head pose estimation, face switching, blink detection, and other tasks have all been effectively accomplished using facial landmarks.

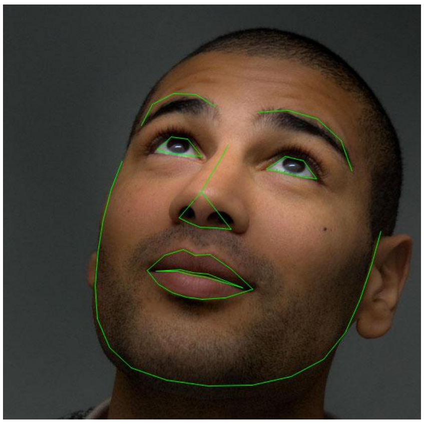

Fig – 3 : In an image, facial landmarks are used to name and identify essential face features.

Face landmark detection is a subset of the shape predicting issue. A shape predictor attempts to pinpoint key points of interest along a shape given an input image (and typically a ROI that describes the item of interest). Our goal in the domain of facial landmarks is to use shape prediction methods to discover essential facial structures just on face.

Detecting facial landmarks is therefore a twostep process:

Step #1: Localize the face in the image.

Step #2: Detect the key facial structures on the face ROI.

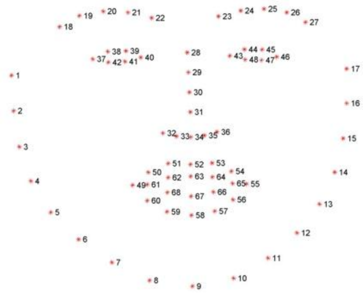

Fig – 4 : The 68 face landmark coordinates from the iBUG 300-W dataset are visualized.

i. EYE ASPECT RATIO

These annotations are from the 68-point iBUG 300-W dataset, which was used to train the dlib face landmark predictor. Starting at the left-corner of the eye (as if staring at the person), and going clockwise around the rest of the region, each eye is represented by 6 (x, y)-coordinates:

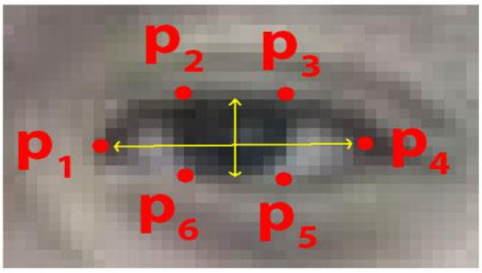

Fig - 5 : The 6 facial landmarks associated with the eye.

ii. MOUTH ASPECT RATIO

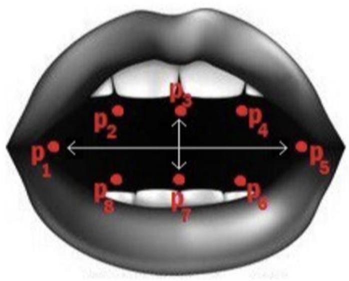

Fig – 6: Facial Landmark associated with Mouth

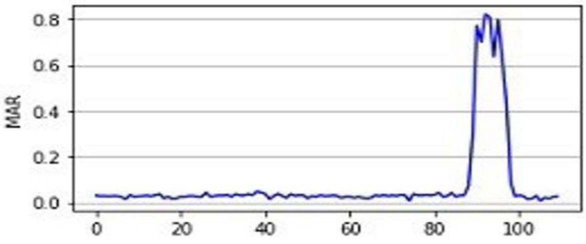

Fig – 7: Graphical Representation of yawning

The graph clearly shows that whenever the mouth is closed, the mouth aspect ratio is nearly zero, as in the first 80 frames. The mouth aspect ratio is higher somewhat when the mouth is partly open. However, in form frame 80, the mouth aspect ratio is much higher, indicating that the mouth is completely open, most likely for yawning.

iii. PnP (For Head Tilt Detection)

In computer vision, an object's pose refers to its orientation and location in relation to a camera. You can alter the attitude by moving the object in relation to the camera or the camera in relation to the item.

A 3D rigid object has only two kinds of motions with respect to a camera.

a. Translation: Translation is the process of moving the camera from its current 3D (X, Y, Z) location to a new 3D location (X', Y', Z'). Translation has three degrees of freedom, as you can see: you can move in the X, Y, or Z directions. Translation is represented by a vector t which is equal to: $(X' - X, Y' - Y, Z' - Z)$

b. Rotation: The camera can also be rotated around the X, Y, and Z axes. As a result, a rotation has three degrees of freedom. Rotation can be represented in a variety of ways. Euler angles (roll, pitch, and yaw), a 3 X 3 rotation matrix, or a rotation direction (i.e. axis) and angle can all be used to express it.

5. OUTCOMES

Fig - 8: Real time view through camera

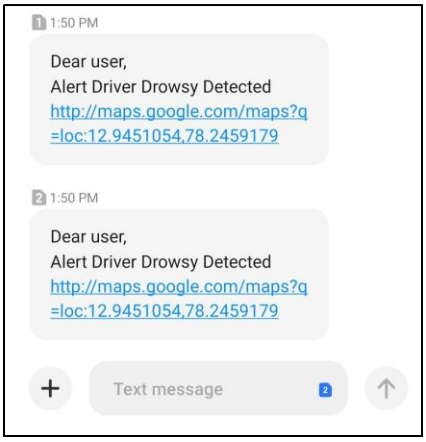

Fig – 9: SMS Alert to Concerned Person

6. CONCLUSION

A framework that spreads and monitors the driver's eye Aspect Ratio, mouth Aspect Ratio, and head movements is designed to detect weariness. The framework utilizes a mixture of layout-based coordinating and highlight-based coordinating to keep the eyes from wandering too far. Framework will almost certainly determine if the driver's eyes are open or closed, and whether the driver is looking ahead, while following. A notice indication will be given as a bell or alarm message when the eyes are closed for an extended period. GPS will detect the live location and update, whenever eyes are closed, continue yawning and head tilted will trigger the system to send SMS alert to provided number ensuring the safety of driver, people walking on the road and other fellow drivers.

REFERENCES

[1] International Organization of Motor Vehicle Manufacturers, "Provisional registrations or sales of new vehicles," http://www.oica.net/wpcontent/uploads/, 2018.

[2] Wards Intelligence, "World vehicles in operation by country, 2013-2017," http://subscribers.wardsintelligence.com/data-browse-world, 2018.

[3] National Highway Traffic Safety Administration, "Traffic safety facts 2016," https://crashstats.nhtsa.dot.gov., 2018.

[4] Attention Technologies, "S.a.m.g-3-steering attention monitor," www.zzzzalert.com, 1999.

[5] Smart Eye, "Smarteye," Available online at: https://smarteye.se/, 2018.

[6] Zhang K, Zhang Z, and Li Z et. al, “Joint face detection and alignment using multitask cascaded convolutional networks,” IEEE Signal Processing Letters, 2016.

[7] M. Danelljan, G. HÃd’ger, F. S. Khan, and M. Felsberg, “Discriminative scale space tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 8, pp. 1561–1575, Aug. 2017.

[8] C. Ma, J. Huang, X. Yang, and M. Yang, “Robust visual tracking via hierarchical convolutional features,” IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1, 2018.

[9] J. Valmadre, L. Bertinetto, J. Henriques, A. Vedaldi, and P. H. S. Torr, “End-to-end representation learning for correlation filter-based tracking,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), July 2017, pp. 5000–5008.

[10] Y. Wu, T. Hassner, K. Kim, G. Medioni, and P. Natarajan, “Facial landmark detection with tweaked convolutional neural networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 3067–3074, Dec. 2018.