Journal of Ubiquitous Computing and Communication Technologies (ISSN: 2582-337X)
www.irojournals.com/jucct/

Enhancing Road Safety: A Driver Fatigue Detection and Behaviour Monitoring System using Advanced Computer Vision Techniques

Ramneet Singh Chadha[1], Jugesh[2], Jasmehar Singh[3]

[1]Associate Director, [2]Project Engineer, [3]Student

[1,2]C-Dac Noida, Uttar Pradesh, India

[3]Shiv Nadar University, Noida, Uttar Pradesh, India

E-mail: [1]rschadha@cdac.in, [2]prajapatijugesh@gmail.com, [3]jasmeharsc@gmail.com

Abstract

Driver drowsiness is a major hazard to road safety, necessitating the development of reliable detection technologies. This study describes a revolutionary driver fatigue detection system that uses cutting-edge computer vision technologies. This system uses the MediaPipe framework for accurate face and hand detection and the Eye Aspect Ratio (EAR) for drowsiness detection. Furthermore, it uses the OpenCV solvePnP function for estimating rotation vector, and converting it to a rotation angle of the head, to check driver attention. By continuously monitoring these indicators, when the system successfully detects any instances of driver tiredness or inattention, it records his or her behavior and delivers notifications to help prevent accidents. This study helps to improve road safety by utilizing cutting-edge computer vision techniques to prevent driver fatigue and boost attentive driving practices.

Keywords: Fatigue detection, Eye aspect ratio, Drowsiness detection, Head pose estimation, Hand detection, Mediapipe

DOI: https://doi.org/10.36548/jucct.2024.2.004
Received: 04.04.2024, received in revised form: 25.04.2024, accepted: 09.05.2024, published: 21.05.2024

1. Introduction

In today's fast-paced environment, driver fatigue and inattention are the major threats to road safety. To solve these issues, an innovative Driver Fatigue Detection System was created, this system uses advanced framework like Mediapipe to detect faces and hands. To assess drowsiness, the system uses the eye aspect ratio (EAR), a measure derived from eye landmarks, by analyzing variations in the EAR over time, it can consistently recognize indicators of tiredness, such as drooping eyelids or extended eye closure. It uses the solvePnP function provided by OpenCV to estimate the driver's head pose by transforming the rotation vector into a rotation angle. This makes it possible to precisely track head tilt and movement, which enables the system to identify unusual behaviors that could be signs of driver distraction or tiredness. Upon detecting signs of drowsiness, one-handed driving, or inattentive behavior, the system sends alert signals to the driver via TCP socket communication. Simultaneously, it records video footage of the driver's behavior, storing it on a server for further analysis and as evidence for vehicle owners. The Driver Fatigue Detection System not only improves road safety but also opens the way for further advancements focused on reducing the effect of driver fatigue.

The methodology employed in this system is significant for several reasons. The ability to process data and generate alerts in real time is crucial for preventing accidents. Immediate feedback to the driver can prompt corrective actions, significantly reducing the risk of accidents caused by fatigue or inattentiveness. The integration of face, eye, and hand detection allows for an assessment of driver behavior. This multi-faceted approach ensures that various forms of distraction and inattention are identified and addressed comprehensively. Recording video footage of detected behaviors provides valuable evidence for vehicle owners or fleet managers to review incidents. This promotes accountability and can be instrumental in post-incident analysis and driver training programs. While the Driver Fatigue Detection System represents a significant improvement in road safety, it is critical to recognize its limits and areas where more research is needed: Poor lighting conditions, occlusions of facial features or hands, and variances in individual behavior can all cause mistakes in fatigue detection. Improving the resilience of the algorithms to reduce false warnings while retaining high detection accuracy remains a major challenge. The use of facial detection technology and the recording of driver behavior creates serious privacy concerns about the acquisition and storage of sensitive data. There is a need for clear norms and guidelines to protect individuals' privacy while also ensuring the system's efficacy in preventing accidents. There is a need for thorough ethical frameworks to govern the development and deployment of these systems, ensuring that driver safety comes first while protecting individual rights and freedoms.

2. Related Work

Fatigue-related accidents continue to be a major concern on the world's roads. Automotive manufacturers and academic researchers have made significant efforts to develop driver monitoring systems for both manually driven and semi-autonomous cars. These systems primarily try to detect driver inattention and fatigue by focusing on factors such as gaze, eye appearance, hand [11], and head movements [8] within the vehicle. Recent advances in computing technology and artificial intelligence have dramatically improved the performance of these systems. Several approaches have been proposed to identify sleepiness signals early, including image-based, biological-based, vehicle-based, and hybrid-based methods [6]. Notably, research has presented novel frameworks for head posture prediction, which use facial landmarks and prominent features to reliably predict head orientation [7]. Furthermore, strong models, such as MFDNet, have been developed to estimate head positions from RGB images, resulting in increased accuracy and reliability [9]. Vision-based hand pose estimate, aided by models such as Google's hand recognition model, provides further insights into driver behavior, allowing for the examination of anticipatory movements during driving maneuvers [10]. Furthermore, driver fatigue monitoring systems that use EAR thresholds and head pose estimate algorithms have been proposed, including real-time alerts and corrective steps to maintain driver attentiveness [12-13]. Recent breakthroughs include principled gaze detection systems that use convolutional neural networks to combine head and eye motions for precise gaze estimates [14].

These improvements collectively help to enhance road safety by addressing driver fatigue and inattention, therefore reducing the likelihood of accidents.

3. Proposed Work

For driver fatigue detection this system harnesses the power of several powerful tools to achieve accurate and reliable results:

MediaPipe: Developed by Google, MediaPipe is a versatile framework that provides pre-trained models adapted to diverse computer vision workloads. Using MediaPipe, the system gains access to advanced features like facial mesh and landmark detection, as well as the ability to recognize hand landmarks. By combining these capabilities, this system can accurately capture facial features and hand motions in real-time, which is critical for assessing the driver's state and detecting indicators of sleepiness or inattention [3].

OpenCV: This open-source computer vision library provides essential features such as video capture, rotation vector calculation, and the production of informative visual cues [4].

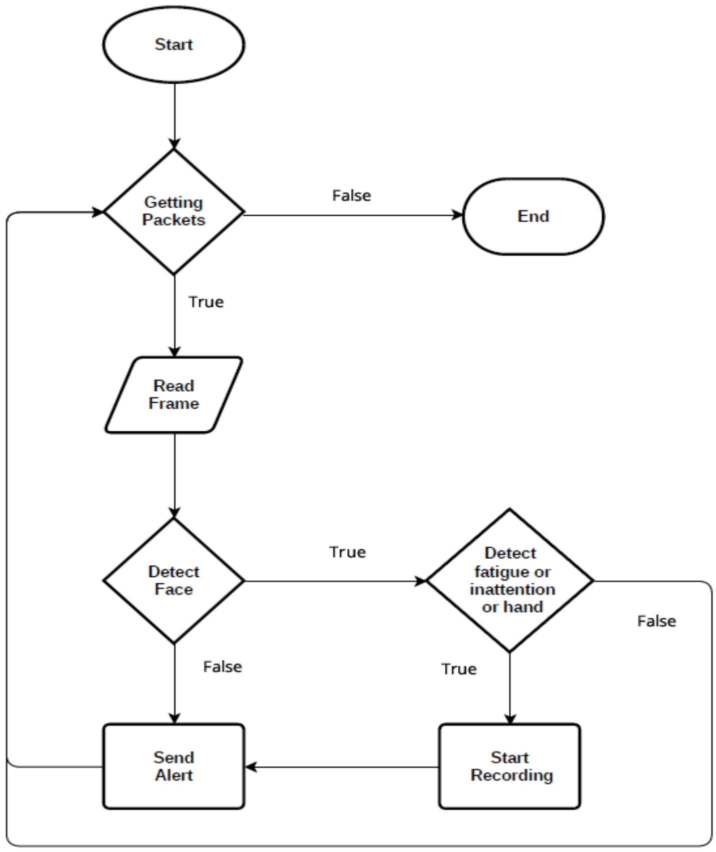

Figure 1. Flow Diagram of Driver Fatigue Detection System.

As shown in Figure 1, the Driver Fatigue Detection System receives packets from the client side if it gets the packets it continues to read the frame using the OpenCV module, this system first detects the face in each frame using 'Mediapipe face solution' (see 3.1) if it detects face it continues to track driver's hand using 'Mediapipe hand solution' (see 3.4), head movement using cv2.solvePnP function (see 3.3) and drowsiness using EAR equation (see

3.2) of the driver, if driver meet any of these conditions this system start to send alert signals and start recording drivers behavior (see 3.5) on server.

3.1 Face Detection

This module detects the driver's face by using a MediaPipe solution, if it detects any face it returns 468 landmark points on the face. MediaPipe discusses the Face Mesh Pipeline is composed of two real-time deep neural network models that work together: A detector that uses the entire image to compute face positions, as well as a 3D face landmark model that uses those sites to forecast the approximate 3D surface using regression. Accurate cropping of the face significantly decreases the requirement for typical data augmentations such as affine transformations, which include rotations, translations, and scale modifications. Instead, it enables the network to devote the majority of its resources to coordinating forecast accuracy [1][3].

3.2 Drowsiness Detection

Continuous driving can be tedious and tiring. Due to a lack of body movement, drivers may feel sleepy while driving. This system determines whether the eyes are closed for an extended period. Because this module focuses on driver drowsiness detection, it only uses landmarks from the eye areas. MediaPipe detects 32 landmarks for eye regions (16 each). To determine whether the eyes are closed, EAR calculation requires only 6 points for each eye.

Equation 1. Eye Aspect Ratio (EAR) Equation, EAR_L for the left eye, EAR_R for the right eye.

$$EAR_L = \frac{\| P_2^L - P_6^L \| + \| P_3^L - P_5^L \|}{2 \| P_1^L - P_4^L \|}$$

$$EAR_R = \frac{\| P_2^R - P_6^R \| + \| P_3^R - P_5^R \|}{2 \| P_1^R - P_4^R \|}$$

Equation 2. Average Eye Aspect Ratio (EAR) Equation

$$Avg.EAR = \frac{EAR_L + EAR_R}{2}$$

For the left eye, the points are [362, 385, 387, 263, 373, 380], and for the right eye, the points are [33, 160, 158, 133, 153, 144]. These landmark points are sequentially labeled as P1, P2, P3, P4, P5, and P6. The average EAR equation returns an average ratio that reflects the level of eye-opening.

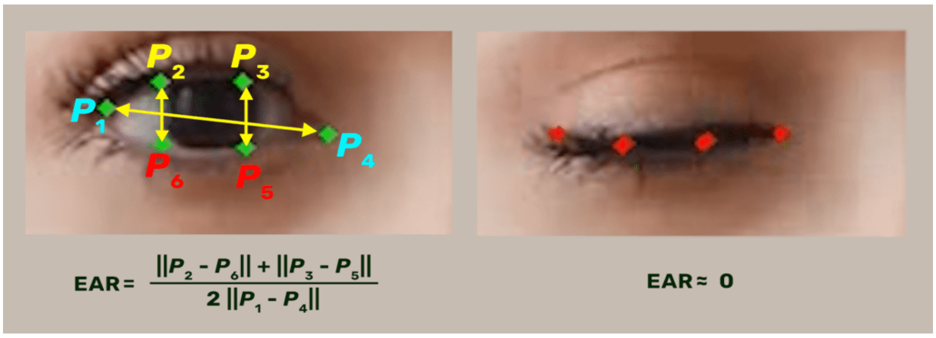

Figure 2. Eye Aspect Ratio (EAR).

In the driver fatigue detection system, various parameters are critical in determining the driver's level of attentiveness. There is a threshold value that is used to establish whether the present Eye Aspect Ratio (EAR) is within an acceptable range, suggesting typical levels of eye openness. The drowsy_frames variable serves as a counter, recording the number of frames in which the EAR falls below the threshold, indicating potential tiredness or inattentiveness. After checking the count of drowsy_frames, the system determines if the driver is asleep or not. To discourage eye blinks, the number of drowsy_frames is reset if the EAR value exceeds the threshold [1].

3.3 Head Pose Estimation

This module tracks the driver's attention and adherence to road safety protocols. It determines whether the driver is focused on the road ahead or distracted by analyzing head movement precisely. This functionality is critical in lowering the dangers associated with driver inattention, which contributes to improved road safety.

The key steps in head pose estimation involve several critical processes. First, MediaPipe is used for facial landmark detection, identifying key features such as the borders of the lips, corners of the eyes, and tip of the nose, which are essential for estimating head posture. To solve the perspective-n-point (PnP) problem and estimate the rotation vector using 3D-to-2D correspondences, the cv2.solvePnP function was used to discover the projection relationship, determining the position of a head in a 3D scene using a series of 2D image points [15]. This function accepts a set of 3D object points and 2D image points, returning a rotation vector that represents the object's position and orientation in 3D space. Transforming these vectors into meaningful angles allows the system to determine the direction the head is facing [2].

3.4 Hand Detection

This module plays an important role in improving road safety by recognizing occasions where the driver's hand is engaged in activities such as using a phone, smoking a cigarette, or eating, which can lead to accidents caused by one-handed driving. It uses the MediaPipe solution's capabilities to precisely detect the presence of hands in the vicinity of the vehicle. Upon detection, the module immediately provides alerts and commences video recording, assuring documentation of any risky activity and permitting early actions to avert accidents.

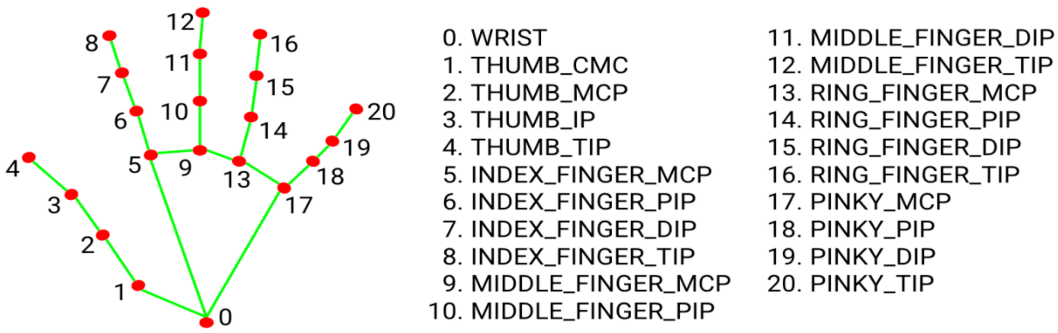

Figure 3. Hand Landmarks.

The hand landmark model bundle recognizes the key point localization of 21 hand-knuckle coordinates inside the identified hand areas. The hand landmark model bundle includes both a palm detection model and a hand landmark detection model. The palm identification model locates hands in the input image, whereas the hand landmarks detection model recognizes specific hand landmarks on the palm [5].

3.5 Recording and Alert System

When the driver meets any of these criteria: Drowsiness when EAR< EAR_threshold, one-handed driving if any hand is detected above steering, or exhibits inattentiveness when head rotation away from a forward-looking position, the system activates an alert and recording mechanism. It starts video recording and saves the video in the file system and its metadata in the PostgreSQL database on a selected server for future reference [16].

Table 1. Recorded Video Metadata Table

Column Name	Data Type
video_id	integer
file_path	varchar
record_date_time	timestamp
video_length(s)	integer
video_size(MB)	double precision
thumbnail	bytea

Table 1 contains metadata for videos recorded by the Driver Fatigue Detection System. Each row in the table represents a distinct video recording. The video_id column contains unique identifiers for each recorded video, it serves as the primary key. The file_path column stores the file path as a string, indicating where the recorded video is saved on the file system. The record_date_time column records the exact date and time when the video was captured. The video_length column indicates the length of the video in seconds. The video_size(MB) column records the size of the video file in megabytes. The thumbnail column stores a binary representation of a thumbnail image extracted from the video. The thumbnail provides a quick visual reference for the content of the video, aiding in easier identification and retrieval of specific recordings.

Furthermore, the system uses TCP (Transmission Control Protocol) socket connectivity for the receiving video stream from driver and quickly sends alert signals to the driver, pushing them to restore focus and attention. This dual capability not only helps the driver be aware but also gives critical proof to the vehicle owner, assisting in post-incident analysis and accountability. Currently, this system uses visual alert as you can see in figure 4,5,6.

4. Experiment and Result

To build the Driver Fatigue Detection system, the Python programming language was used along with the MediaPipe library, OpenCV library, and several other built-in libraries. The system was developed and tested on a local machine running Windows OS. The experiment was designed to assess the performance of the system in identifying head position, drowsiness, and hand motions in an indoor setting with consistent lighting conditions. A webcam was utilized to record live video footage of the driver's face and hands. Participants were asked to sit in a stationary seat and simulate driving while being monitored by the Driver Fatigue Detection System. Several scenarios were replicated, including periods of concentration, occasions of drowsiness, and the use of hands for other tasks such as phone calls or meals. Any observed instances of inattention, drowsiness, or hand engagement in another task, resulted in a visual alert signal. The system simultaneously records video of the driver's inattentive behavior in the file system and saves its metadata on the server, see Table 1.

Figure 4. Head Pose Estimation

The Figure 4 showed that the system can accurately predict the driver's head pose in real-time. The device accurately recorded head motions and rotations, providing vital information about the driver's behavior.

Figure 5. Drowsiness Detection

The Figure 5 depicts that Driver Fatigue Detection System correctly identified indicators of drowsiness based on changes in the driver's eye movements using Eye Aspect Ratio (EAR). The alert signal was triggered when the EAR fell below a predefined threshold, suggesting potential drowsiness.

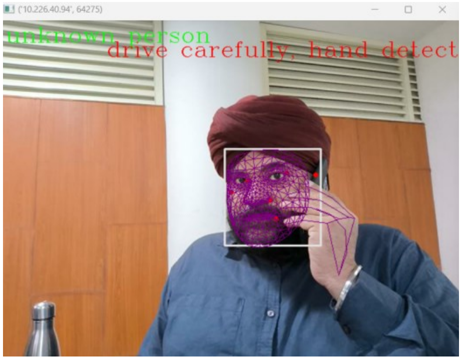

Figure 6. Hand Detection

Throughout the experiment, the system reliably recognized hand motions as shown in Figure 6 and other activities such as phone use or eating. When the system spotted such behaviors, it generated alerts that highlighted possible distractions for the driver.

Overall, the trial proved that the Driver Fatigue Detection System is effective at monitoring driver behavior and recognizing instances of inattention in an indoor context. The acquired video footage and alert signals give vital information for further study and optimization of the system's functioning.

5. Conclusion

When the system detects fatigue or inattentive driving behavior, it immediately records the driver's actions and issues an alarm to regain attention, averting potential accidents and protecting the safety of both the driver and other road users. The Driver Fatigue Detection System is a critical tool for improving road safety by reducing the dangers associated with driver fatigue and distraction. Furthermore, the system's future scope looks optimistic, with prospective technological breakthroughs paving the way for even more powerful and complex solutions. Sensor technology and computational power advancements may allow for more compact and cost-effective implementations, making the system available to a broader range of cars and drivers. In conclusion, while the Driver Fatigue Detection System has significant promise for improving road safety, overcoming its limits will necessitate interdisciplinary research efforts spanning computer vision, machine learning, human-computer interaction, and ethics. By resolving these issues, subsequent generations of the system can increase its effectiveness and dependability, resulting in a safer driving environment.

References

[1] Vaibhav Singh (2023, May 9). Driver Drowsiness Detection Using Mediapipe In Python. Accessed on (2024, January 2) https://learnopencv.com/driver-drowsiness-detection-using-mediapipe-in-python/

[2] R, J. (2023, September 22). Real-Time Head Pose Estimation FaceMesh with MediaPipe and OpenCV: A Comprehensive Guide. Medium. Accessed on (2024, January . https://medium.com/@jaykumaran2217/real-time-head-pose-estimation-facemesh-with-mediapipe-and-opencv-a-comprehensive-guide-b63a2f40b7c6

[3] Lugaresi, Camillo, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang et al. "Mediapipe: A framework for building perception pipelines." arXiv preprint arXiv:1906.08172 (2019).

[4] I. Culjak, D. Abram, T. Pribanic, H. Dzapo and M. Cifrek, "A brief introduction to OpenCV," 2012 Proceedings of the 35th International Convention MIPRO, Opatija, Croatia, 2012, pp. 1725-1730.

[5] Hand landmarks detection guide for Python. (2023, September 03). Google for Developers. Accessed on (2024, January 5) https://developers.google.com/mediapipe/solutions/vision/hand_landmarker/python

[6] Albadawi, Yaman, Maen Takruri, and Mohammed Awad. "A review of recent developments in driver drowsiness detection systems." Sensors 22, no. 5 (2022): 2069.

[7] Barros, Jilliam María Díaz, Frederic Garcia, Bruno Mirbach, Kiran Varanasi, and Didier Stricker. "Combined Framework for Real-time Head Pose Estimation using Facial Landmark Detection and Salient Feature Tracking." In VISIGRAPP (5: VISAPP), pp. 123-133. 2018.

[8] Barros, Jilliam Maria Diaz, Bruno Mirbach, Frederic Garcia, Kiran Varanasi, and Didier Stricker. "Fusion of keypoint tracking and facial landmark detection for real-time head pose estimation." In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 2028-2037. IEEE, 2018.

[9] Liu, Hai, Shuai Fang, Zhaoli Zhang, Duantengchuan Li, Ke Lin, and Jiazhang Wang. "MFDNet: Collaborative poses perception and matrix Fisher distribution for head pose estimation." IEEE Transactions on Multimedia 24 (2021): 2449-2460.

[10] Zhang, Fan, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. "Mediapipe hands: On-device real-time hand tracking." arXiv preprint arXiv:2006.10214 (2020).

[11] Das, Nikhil, Eshed Ohn-Bar, and Mohan M. Trivedi. "On performance evaluation of driver hand detection algorithms: Challenges, dataset, and metrics." In 2015 IEEE 18th international conference on intelligent transportation systems, pp. 2953-2958. IEEE, 2015.

[12] Biswal, Anil Kumar, Debabrata Singh, Binod Kumar Pattanayak, Debabrata Samanta, and Ming-Hour Yang. "IoT-based smart alert system for drowsy driver detection." Wireless communications and mobile computing 2021 (2021): 1-13.

[13] Zhang, Meng, and Fumin Zhang. "A Driver Fatigue Detection Method Based on Eye Aspect Ratio Compensated by Head Pose Estimation." In 2023 8th International Conference on Signal and Image Processing (ICSIP), pp. 161-165. IEEE, 2023.

[14] Jha, S., Al-Dhahir, N., & Busso, C. (2023). Driver Visual Attention Estimation Using Head Pose and Eye Appearance Information. IEEE Open Journal of Intelligent Transportation Systems, 4, 216–231.

[15] Z. Zhang, "A flexible new technique for camera calibration," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, Nov. 2000,

[16] Stonebraker, M., Rowe, L., & Hirohama, M. (1990, March). The implementation of POSTGRES. IEEE Transactions on Knowledge and Data Engineering, 2(1), 125–142.

Author's biography

Ramneet Singh Chadha -Ramneet Singh Chadha is currently working as Associate Director in C-DAC, Noida. He has more than 25 years of experience in IT expertise. He has a Master's in Management (IT) and M.S. (Software Systems) from BITS Pilani. He was a College topper while pursuing his Bachelor in Computer Engineering from Nagpur University. He has vast experience in the Health care Domain and Transit Domain. He has experience in developing and implementing National/Statewide Hospital projects and developed the NCMC Standard for MOHUA in the transit domain. He is extensively working in ITMS (Integrated Transit Management System), and Smart Card Based ID and Access Control System for MHA (Ministry of Home Affair). He has also experience in academia as well as industry.

Jugesh -Jugesh is a Project Engineer in the Department of Embedded Systems at C-DAC, Noida. He completed his master's in Computer Application from C-DAC, Noida in 2023. He Graduated in Software Development from GNDIT Rohini affiliated with GGSIPU in Delhi, India. He has great inspiration and deep interest in Research work in the field of Artificial Intelligence, Computer Vision and Machine Learning.

Jasmehar Singh -Jasmehar Singh is a 2nd-year student pursuing Computer Science from Shiv Nadar University with a keen interest in Data science, Computer Vision and Artificial Intelligence. He is passionate about exploring and solving the problems in the transit domain. He has a great interest in Research work in the fields of AI, ML, and cyber security.