Grenze International Journal of Engineering and Technology, June Issue

Drowsiness Detection using DCCNN Algorithm

Harika G[1], Swathi N[2], Sai Nikhil M[3], Charan Teja J[4], Praveen Kumar. K[5], and Md. Firose Shaik[6]
[1-6]VR Siddhartha Engineering College/EIE, Vijayawada, India
Email: harikagude2004@gmail.com, swathinadipineni@vrsiddhartha.ac.in, nikhilreddy19142@gmail.com, tejac212@gmail.com, kunchalapraveen100@gmail.com, firose@vrsiddhartha.ac.in

***Abstract*— The rising global prevalence of vehicles has led to an escalation in traffic accidents, emerging as a major contributor to human fatalities. This research focuses on improving traffic safety by addressing the issue of drowsy driving. It suggests an algorithm for detecting driver drowsiness in real-time, considering the unique differences among individual drivers. The algorithm uses a Deep Cascaded Convolutional Neural Network (DCCNN) to detect the driver's face, including landmarks with the help of the Dlib toolkit. A novel metric termed "Eye Aspect Ratio (EAR)" is introduced to assess drowsiness by analyzing the landmarks of the driver's eyes. The algorithm consists of two modules: offline training, where a fatigue state classifier is trained, and online monitoring, where the classifier is applied to detect drowsiness quickly from driver images. It employs a variable determined by the frequency of drowsy frames per unit of time for drowsiness evaluation. Comparative experiments indicate that this algorithm surpasses current drowsiness detection methods in terms of both accuracy and speed. The findings from this research have the potential to enhance intelligent transportation systems, improve driver safety, and mitigate losses associated with drowsy driving.**

***Index Terms*— DCCNN, Dlib tool kit, Eye Aspect ratio, Drowsiness detection, Driver safety.**

I. INTRODUCTION

Detecting drowsiness is crucial for road safety, especially in driver monitoring systems. Deep Cascade Convolutional Neural Networks (CNNs) have become a powerful tool in this field. They automatically identify signs of drowsiness in real-time video feeds, helping prevent accidents caused by fatigued or inattentive drivers. In this introduction, we'll explore the application of Deep Cascade CNNs in drowsiness detection, highlighting their potential to enhance road safety, reduce accidents, and even save lives [1]. We'll dive into how this technology works, its advantages over traditional methods, and its importance in proactive drowsiness detection, applicable in driver assistance systems and commercial vehicle fleets.

A. Causes of Drowsiness and their Prevention

a) Microsleep: Microsleeps are brief episodes of sleep lasting mere seconds, distinct from the longer sleep periods of minutes or hours. These symptoms commonly appear following insufficient sleep, impacting individuals dealing with sleep disorders such as shift work disorder or obstructive sleep apnea. However, even those without sleep disorders can encounter microsleeps, even following just one night of limited sleep or no sleep deprivation at all. These fleeting moments of sleep frequently occur during monotonous activities, such as driving on deserted highways.

Grenze ID: 01.GIJET.10.2.143
© Grenze Scientific Society, 2024

b) Symptoms of Microsleep: Microsleep is often characterized by the common symptom of partially or fully closing your eyes, though it can also occur with your eyes open. Another typical indicator of microsleep is nodding your head. Rather than realizing that you've briefly drifted off to sleep, you may perceive it as a momentary lapse in attention to your surroundings. During a microsleep episode, individuals exhibit a diminished responsiveness to external stimuli, including auditory and visual cues. Studies have revealed that in the moments leading up to a microsleep episode, people tend to exhibit slower eye movements. However, you may not be consciously aware of these slower eye movements or the drooping of your eyelids [2]. Pupil dilation is another noteworthy sign that may suggest the occurrence of microsleep.

c) Risk of microsleep: The principal dangers associated with microsleep are the potential for accidents during activities such as driving, operating heavy machinery, performing surgery, or engaging in similarly critical tasks. Research conducted on simulated driving revealed that microsleep episodes were linked to a decline in driving performance. It is worth noting that experiencing microsleep does not inherently cause physical harm, as illustrated in "Fig.1". In cases where microsleep occurs in a safe environment, where a momentary lapse of attention does not have significant consequences, the occurrence of microsleep should not pose a concern.

Figure 1. Motor vehicle accident caused by drowsy driving

d) Prevention of microsleep: At present, there is no universally endorsed treatment for microsleep, and microsleep itself is not classified as a distinct sleep disorder nor included in the diagnostic criteria for any sleep-related condition. Given that microsleep is closely linked to sleep deprivation and drowsiness, ensuring you receive sufficient sleep may offer a potential remedy. Wakefulness-promoting medications might be considered to mitigate the occurrence of microsleep episodes; however, their usage should be approached cautiously and only under the guidance of a medical professional.

II. PROBLEM DESCRIPTION

Drowsiness detection while operating vehicles or engaging in activities requiring sustained attention is a crucial concern for safety. Drowsiness can lead to reduced reaction times, impaired decision-making, and an increased risk of accidents. To tackle this concern, the challenge of detecting drowsiness is addressed through the application of a DCCNN algorithm. The key contributions of this initiative can be outlined as follows: we enhanced the DCCNN for driver face detection, leading to improved performance; additionally, we introduced a parameter EAR based on the Dlib toolkit to evaluate whether the driver's eyes are open or closed. The initial phase constitutes a Fully Convolutional Network (FCN). Unlike a CNN, an FCN lacks a dense layer in its architecture. This Proposed Network is employed to acquire candidate windows along with their bounding box regression vectors.

There are five key facial landmarks: the left eye, right eye, nose, left mouth corner, and right mouth corner.

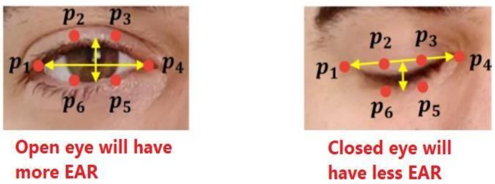

Figure 2. Calculation of EAR

Dlib is a modern C++ toolkit equipped with machine learning algorithms and tools designed for solving real-world problems in C++. Widely applied in both industry and academia, it finds utility in diverse domains, including robotics, embedded devices, mobile phones, and large-scale high-performance computing environments. The EAR is a scalar value specifically responsive to the opening and closing of eyes. To discern between open and closed eye states, an EAR threshold is utilized, illustrating the temporal progression required to calculate a typical EAR value for a single blink, as depicted in "Fig. 2". During the blinking process, notable fluctuations in the EAR value are observed.

III. LITERATURE SURVEY

In recent years, research into methods for detecting driving drowsiness based on driver behavior has gained significant attention due to their non-invasive and cost-effective nature. Nevertheless, the effectiveness of these algorithms can be constrained by the capabilities of face detection and fatigue assessment technologies, especially in dynamic and unpredictable environments.

A. AdaBoost-based Face Detector

In reference to [3], when the driver is directly facing the camera, a face detector based on AdaBoost is employed to determine the position and scale of the face. Subsequently, an Active Shape Model (ASM) algorithm is utilized to precisely pinpoint the location of the human eyes in front-view images. Test results reveal significant variations in eye movements corresponding to different levels of drowsiness. In a fully awake state, eye movements are characterized by low-frequency, high-speed motions with larger mean amplitudes. However, as drowsiness sets in, the mean amplitude of eye movements decreases, accompanied by a significant increase in frequency. As drowsiness escalates, the mean amplitude of eye movements further diminishes, and prolonged eye closures become more pronounced in a highly drowsy state.

B. Light-intensive System

In the study conducted by Malla and colleagues [4], they have devised a system that is resilient to variations in lighting conditions. Their approach involved employing the Haar algorithm for object detection and utilizing a face classifier implemented within the OpenCV libraries. From the facial region, they extracted eye regions based on anthropometric factors and proceed to detect the eyelids to assess the degree of eye closure. On the other hand, Vitiable and associates [5] have developed a drowsiness detection system using an infrared camera. Their system leverages the phenomenon of bright pupils to create an algorithm capable of detecting and tracking the driver's eyes. When signs of drowsiness are identified, the system promptly alerts the driver with an alarm message.

C. Otsu Thresholding System

In their work, Bhowmick et al. [6] utilized Otsu thresholding to extract the facial region. The eye's localization is achieved by identifying facial landmarks such as the eyebrow and a potential face center. They employed morphological operations and K-means for precise eye segmentation. Subsequently, a set of shape features is computed and then trained using a non-linear Support Vector Machine (SVM) to determine the eye's condition. As for Hong and colleagues [7], They developed a real-time system for drowsiness detection by analyzing eye states. Facial region detection is initiated using the optimized Jones and Viola method [8], followed by horizontal projection to obtain the eye area. The eye state is then determined using a novel complexity function with a dynamic threshold.

IV. PROPOSED METHOD

To effectively identify driver drowsiness, we have chosen to implement an advanced face detection algorithm known as the DCCNN. We employed the Dlib toolkit to pinpoint facial landmarks, and we have utilized the EAR to gauge the level of drowsiness. Additionally, we also considered a critical factor: the unique characteristics of each driver. For face detection, the DCCNN algorithm is employed, which comprises three internal layers crucial for precise image recognition. The DCCNN architecture closely resembles that of the Multi-Task Convolution Neural Network (MTCNN), with one key distinction: MTCNN utilizes OpenCV to extract facial landmarks, while DCCNN relies on the Dlib toolkit to obtain these landmarks on the face. Dlib toolkit, a C++ library, allows us to access 68 facial landmarks, which is essential for accurately calculating the EAR. This choice of Dlib over OpenCV is primarily motivated by the need for a more comprehensive set of facial landmarks, as OpenCV provides only five. To achieve a reliable drowsiness detection mechanism, we leveraged the Dlib toolkit to obtain facial landmarks, which, in turn, enables us to compute the EAR. Once EAR values are obtained, apply the P80 algorithm to detect signs of drowsiness. When drowsiness is detected, an alert alarm is activated to promptly notify the driver.

V. DESIGN FLOW OF THE PROPOSED SOLUTION

In this work there are two modules. They are ‘Offline training’ and ‘Online monitoring’ and their design flow is shown in “Fig. 3” and “Fig. 4” respectively.

Offline Training

- Step 1: At first, capture the photo.
- Step 2: If photo is captured, then face will be detected. Else print “Picture not captured”.
- Step 3: If face detected, then obtain landmarks. Else print “Face not detected”.
- Step 4: After landmarks obtained, calculate EAR ratio and based on EAR calculate P80.

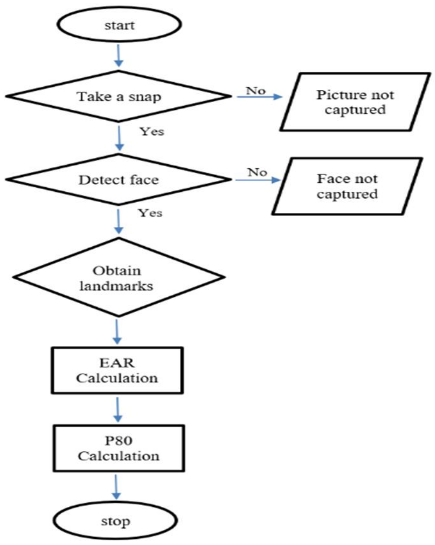

Figure 3. Design flow of offline training

Online Monitoring

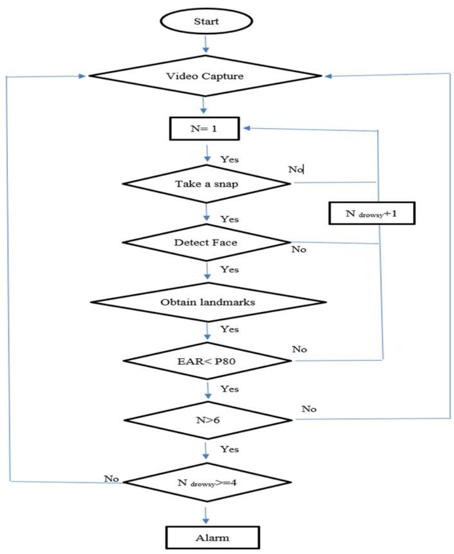

Figure 4. Design flow of online monitoring

- Step 1: In online monitoring detect face, obtain land marks and calculate EAR.
- Step 2: If face is not detected or EAR<P80, increase number of drowsy frames. Else check N>6 or not.
- Step 3: If N>6 and number of drowsy frames greater than or equal to 4 then alarm rings. Else the process

repeats until the alarm rings.

VI. SOFTWARE IMPLEMENTATION

A. OpenCV

OpenCV, short for the OpenCV, a freely available and open-source software library, is specifically crafted for applications in computer vision and machine learning. Its primary purpose is to establish a unified foundation for computer vision projects and promote the integration of machine perception in commercial products [9].

B. Play sound Module

The "play sound" module consists of a single function with the same name, "play sound." This function takes one mandatory argument: the path to the sound file you want to play. This path can point to either a local file or a URL. Additionally, there's an optional second argument called "block," which defaults to "True." When set to "False," this argument enables the function to run asynchronously. On Windows, the module utilizes the "Win MM" library for sound handling [10]. It has been tested and confirmed to work with WAVE and MP3 file formats. Although other file formats may function properly, these two are guaranteed to be supported.

C. MTCNN Algorithm

A Python 3.4+ implementation of the MTCNN for face detection, employing TensorFlow and Kera's. This implementation has been built from scratch, with reference to David Sandberg's MTCNN implementation from Face Net [11].

D. Dlib toolkit

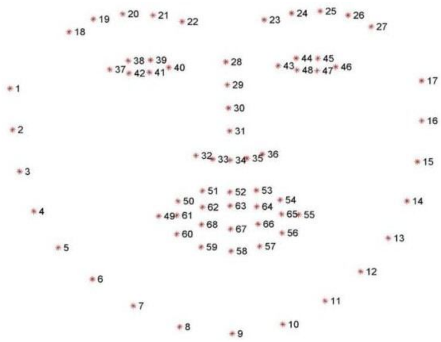

Figure 5. Facial landmarks using Dlib toolkit

Dlib is a contemporary C++ framework equipped with machine learning algorithm and software development tool. Dlib allows us to access 68 facial landmarks, as shown in "Fig. 5", which is essential for accurately calculating the EAR. It empowers developers to tackle real-world challenges in various domains, including but not limited to robotics, embedded systems, mobile devices, and high-performance computing environments. making it a popular choice in both industry and academia [12].

VII. RESULTS AND DISCUSSION

Offline Training

In the offline training phase, the system analyses two input images. These images depict the driver's eyes in distinct states: open and closed. The primary objective during training is to extract EAR values. EAR is a crucial metric for assessing the eye state based on geometric ratios. The first input image captures the open-eye scenario, as illustrated in "Fig. 6". Conversely, the second image represents the closed-eye condition, as shown in "Fig. 7". The training process aims to discern patterns and features related to eye states. EAR ratios are computed for both open and closed frames during this training. Additionally, the extraction process includes the determination of P80 metrics. The collected data from this training is fundamental for subsequent real-time drowsiness detection.

Online Monitoring

In the online monitoring module, a sequence of 6 frames serves as input. These frames are depicted in "Fig. 10," "Fig. 11," "Fig. 12," "Fig. 13," "Fig. 14," and "Fig. 15." The primary goal is to assess whether the driver is experiencing drowsiness. The system continuously analyses the input frames for signs of drowsiness. Upon detecting drowsiness, an alarm is promptly triggered. If no drowsiness is detected, the monitoring process continues. The online module operates iteratively until drowsiness is identified. This iterative process ensures real-time assessment of the driver's alertness.

Figure 6. Closed Frame

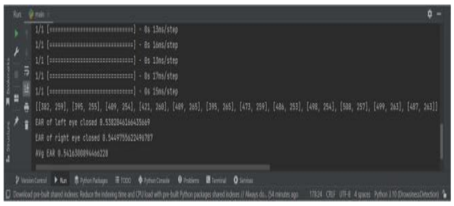

Figure 7. EAR of closed frame

Figure 8. Open Frame

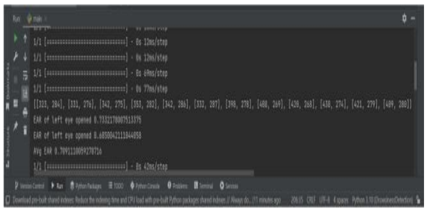

Figure 9. EAR of open frame

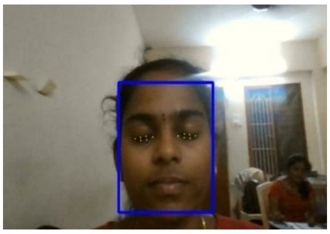

Figure 10. Frame-1

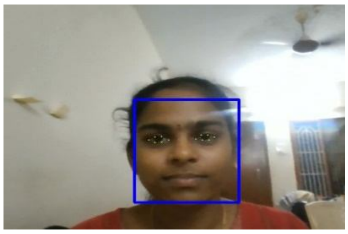

Figure 11. Frame-2

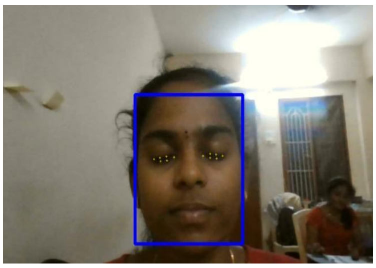

Figure 12. Frame-3

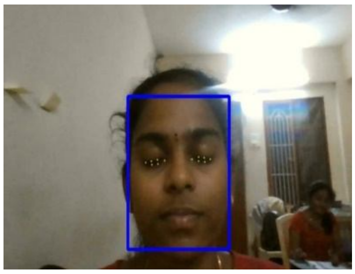

Figure 13. Frame-4

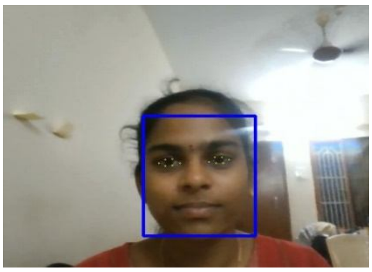

Figure 14. Frame-5

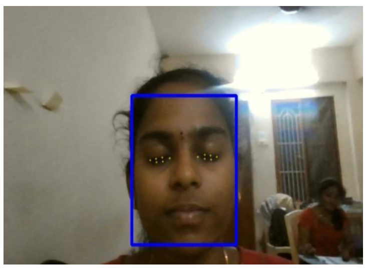

Figure 15. Frame-6

By the above input images, we can observe that in “Fig. 10”, “Fig. 12”, “Fig. 14”, and “Fig. 15” the person is clearly drowsy. Now, calculate PERCLOS by using the formula

$$\text{PERCLOS} = Ndrowsy/Ntotal$$

Here Ndrowsy = 4 and Ntotal = 6. Then PERCLOS is 0.666666, which is greater than 0.4 and so, an alarm will be raised.

VIII. Conclusions

This work introduced an innovative algorithm for detecting driving drowsiness while considering individual differences. Initially, we've developed a convolutional neural network model called DCCNN. This model eliminates the requirement for manual feature extraction in conventional face detection methods, allowing us to extract a driver's face from live video. We've rigorously evaluated the model's performance through both qualitative and quantitative assessments. To assess the driver's eye state, we introduced a new metric called EAR, which relies on the Dlib toolkit. In comparison to conventional methods, EAR proves to be more consistent, credit goes to the Cascaded Pose Regression algorithm. Our experiments revealed a robust association between EAR and the size of a driver's eyes, validating the soundness of approach. Considering the variations among drivers, we've developed an offline training module and an online monitoring module. These modules use unique classifier based P80 to individualize the assessment of a driver's eye state during driving. Our experimental results consistently outperform current state-of-the-art methods while maintaining real-time performance.

References

[1] National Traffic Safety Administration, "Drowsy Driving," 25 December 2017. [Online]. Available: https://www.nhtsa.gov/riskydriving/drowsy-driving.

[2] Sleep foundation, "Microsleep," One care media company, 6 January 2023. [Online]. Available: https://www.sleepfoundation.org/how-sleep-works/microsleep.

[3] Z. Wei, B. Cheng and Y. Lin, "Driver drowsiness recognition based on computer vision technology," IEEE, vol. 17, no. 3, pp. 354-362, 2012.

[4] Malla AM, Davidson PR, Bones PJ, Green R, Jones RD. Automated video-based measurement of eye closure for detecting behavioral microsleep. Annu Int Conf IEEE Eng Med Biol Soc. 2010; 2010:6741-4. doi: 10.1109/IEMBS.2010.5626013. PMID: 21095829.

[5] S. Vitabile, A. De Paola and F. Sorbello, "Bright Pupil Detection in an Embedded, Real-Time Drowsiness Monitoring System," 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, WA, Australia, 2010, pp. 661-668, doi: 10.1109/AINA.2010.151.

[6] B. Bhowmick and K. S. Chidanand Kumar, "Detection and classification of eye state in IR camera for driver drowsiness identification," 2009 IEEE International Conference on Signal and Image Processing Applications, Kuala Lumpur, 2009, pp. 340-345, doi: 10.1109/ICSIPA.2009.5478674.

[7] T. Hong, H. Qin and Q. Sun, "An Improved Real Time Eye State Identification System in Driver Drowsiness Detection," 2007 IEEE International Conference on Control and Automation, Guangzhou, China, 2007, pp. 1449-1453, doi: 10.1109/ICCA.2007.4376601.

[8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[9] W. Kongcharoen, S. Nuchitprasitchai, Y. Nilsiam and J. M. Pearce, "Real-Time Eye State Detection System for Driver Drowsiness Using Convolutional Neural Network," 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 2020, pp. 551-554, doi: 10.1109/ECTI-CON49241.2020.9158265.

[10] OpenCv, "Opencv Documentation," [Online]. Available: https://opencv.org/about/.

[11] Pypi, "Playsound-pypi," [Online]. Available: https://pypi.org/project/playsound/.

[12] World Health Oraganisation, "World health Oraganisation," 20 June 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.