Driver Fatigue Detection Method Based on Eye States With Pupil and Iris Segmentation

PSD 1 report submitted in partial fulfillment of the
Requirements for the degree of

Master of Engineering
ME (Big Data Analytics)

by

Manthana H K
231058012

MANIPAL
ACADEMY *of* HIGHER EDUCATION
(Deemed to be University under Section 3 of the UGC Act, 1956)

INSPIRED BY LIFE

MANIPAL SCHOOL OF INFORMATION SCIENCES
(A Constituent unit of MAHE, Manipal)

Acknowledgement

It is great pleasure to thank the people behind the success of this PSD 1 activity. I owe my deepest gratitude to all those who guided, inspired, and helped me to complete this project.

I would like to take this opportunity to express my gratitude and heartiest thanks to my panel members, **Prof. Arockiaraj S**, Assistant Professor and **Prof. Prathviraj N,** Assistant Professor Manipal School of Information Sciences for his inspirational support and guidance throughout my project period.

My heartful gratitude to **Dr Keerthana Prasad** Professor & Director Manipal School of Information Sciences for his full support and encouragement during the PSD 1 activity.

I would like to thank all sources mentioned in the references.

LIST OF TABLES

LIST OF FIGURES

Contents

Abstract

Fatigue driving has become one of the most common causes for traffic accidents. In this paper, authors have proposed an effective fatigue detection method based on eye status with pupil and iris segmentation. The segmented feature map can guide the detection to focus on pupil and iris. A streamlined network, consisting of a segmentation network and a decision network, is designed, which greatly improves the accuracy and generalization of eye openness estimation. Specifically, the segmentation network that uses light U-Net structure performs a pixel-level classification on the eye images, which can accurately extract pupil and iris features from the video's images. Then, the extracted feature map is used to guide the decision network to estimate eye openness. Finally, the detection method is tested by the National Tsing Hua University Drowsy Driver Detection (NTHU-DDD) Video Dataset and the precision of fatigue detection achieves 96.72%. Experimental results demonstrate that the proposed method can accurately detect the driver fatigue in-time and possesses superior accuracy over the state-of-the-art techniques.

1. Chapter 1
Introduction

1.1 Introduction

Driving for a long time will cause driver fatigue. And the driver will have a phenomenon of drowsiness, which easily leads to a traffic accident. In the 2018 Global Road Safety Status Report, World Health Organization (WHO) noted that 1.35 million people died in road traffic accidents every year. According to the different fatigue performances, there are the following three kinds of detection methods: (1) fatigue detection method based on the physiological signals, which is to judge fatigue according to the oxygen level of brain tissues and muscle tissues, (2) fatigue detection method based on the vehicle behavior, which monitors the variations of steering wheel angle, lane position, speed, acceleration, and braking to predict the driver fatigue and (3) fatigue detection method based on the drivers' facial features, which monitors the facial expressions to judge the driver's fatigue state because the frequency of yawning and blinking increases when the driver is fatigued.

Figure 1: General framework for eye status detection

Third kind of method, which only depend on vision analysis to monitor the driver's status, is non-intrusive and does not interfere with the driving behavior. PERCLOS (percentage of eyelid closure over the pupil) was an effective indicator for fatigue detection. In order to accurately calculate PERCLOS, they have judged the eye status in real time.As shown in Fig.1, in most of them the eye images are directly taken as the input, and then the Convolutional Neural Networks (CNN) is used to judge to the eye status. The paper have proposed a segmentation based eye status detection model (SESDM). The model consists of a segmentation network and a decision network. It not only improves the accuracy of the network, but also improves the generalization of the network.

.

Accurate eye status detection method is very important for fatigue detection. There are various methods for estimating the eye status, which are based on shape recognition, template matching, traditional machine learning and deep learning. Shape-based approaches aim to recognize the eye states according to the geometric relationships or circular shape of visible iris. Template matching is another simple and direct method of eye status detection. However, using shape recognition or template matching, the different parameters need to be set for different individuals. In a real environment, varying lighting conditions will lead to increased difficulty in eye status detection. Template matching, shape recognition, and traditional learning method have poor detection performance in the complex environment. In summary, shape recognition and the template matching methods have the problem of insufficient generalization. Traditional machine learning methods can divide the eyes into open, closed, and half-open states. However, it is fuzzy for the method to judge the process from closed to open. Hence, to accurately characterize the driver drowsiness based on the eye images, it is desirable to measure the eye-openness continuously.

1.2 Literature Review

Rateb Jabbara et.al, [1] proposed a drowsiness detection system based on multilayers perceptron classifiers. It is specifically designed for embedded systems such as Android mobile. The role of the system is to detect facial landmark from images and deliver the obtained data to the trained model to identify the driver's state. The purpose of the method is to reduce the model's size considering that current applications cannot be used in embedded systems due to their limited calculation and storage capacity. According to the experimental results, the size of the used model is small while having the accuracy rate of 81%. Hence, it can be integrated into advanced driver-assistance systems, the Driver drowsiness detection system, and mobile applications. However, there is still space for the performance improvement. The further work will focus on detecting the distraction and yawning of the driver. Experimental results show that the accuracy rate is almost 81%. When it comes to the model size, its complexity and storage, there is a significant reduction when the face landmark coordination is used to detect the driver's drowsiness. The maximum size of the developed model is 100 KB.

Cyun-Yi Lin et al, [2] proposed the machine learning and gradient statistics based technologies, the driver drowsiness detection is proposed for the real-time application. The proposed design uses gray-scale images without any color information, and it works effectively in daytime and nighttime. In experimental results, the average processing frame rates are up to 245 fps by a PC with 2.59GHz operational frequency. The average detection rate of eye closure is 91.49% when the driver wears glasses, and the corresponding detection rate is 95% when the driver does not wear glasses. the average detection accuracy of open/closed eyes by the proposed design is 91.49%.The proposed a system by recognizing the accurate eye position, the effective eye detection methodology is proposed to enhance the accuracy of fatigue detections, and the machine learning and image gradients based schemes are utilized for the driver drowsiness detection design.

Wanghua Deng and Ruoxue Wu et al, [3] proposed a novel system for evaluating the driver's level of fatigue based on face tracking and facial key point detection. They have designed a new algorithm and propose the MC-KCF algorithm to track the driver's face using CNN and MTCNN to improve the original KCF algorithm. They have defined the facial regions of detection based on facial key points. Moreover, they have introduced a new evaluation method for drowsiness based on the states of the eyes and mouth. Therefore, DriCare is almost a real-time system as it has a high operation speed. From the experimental results, DriCare is applicable to different circumstances and can offer stable performance. When a driver is in a state of fatigue, the facial expressions, e.g., the frequency of blinking and yawning, are different from those in the normal state. In this paper, the authors have proposed a system called DriCare, which detects the drivers' fatigue status, such as yawning, blinking,and duration of eye closure, using video images, without equipping their bodies with devices. Owing to the shortcomings of previous algorithms, they have introduced a new face-tracking algorithm to improve the tracking accuracy. Further, they have designed a new detection method for facial regions based on 68 key points. Then they have used these facial regions to evaluate the drivers' state. By combining the features of the eyes and mouth, DriCare can alert the driver using a fatigue warning. The experimental results showed that DriCare achieved around 92% accuracy.

Burcu Kir Savaş and Yaşar Becerİklİ et al, [4] proposed a Multi-tasking Convulational Neural Network (ConNN) model is proposed to detect driver drowsiness/fatigue. Eye and mouth characteristics are utilized for driver's behavior model. Changes to these characteristics are used to monitor driver fatigue. With the proposed Multi-task ConNN model, unlike the studies in the literature, both mouth and eye information are classified into a single model at the same time. Driver fatigue is determined by calculating eyes' closure duration/Percentage of eye closure (PERCLOS) and yawning frequency/frequency of mouth (FOM). In this study, the fatigue degree of the driver is divided into 3 classes. The proposed model achieved 98.81% fatigue detection on YawdDD and NthuDDD dataset. The success of the model is presented comparatively. The proposed system can model the interactive relationship between eye, mouth and sub-states. Fatigue at a predetermined time point is considered a factor for fatigue at the present time point, and time varies according to the behavior of individuals.

Md. Tanvir Ahammed Dipu et al, [5] the proposed system for detecting driver drowsiness is used in the Convolutional Neural Network (CNN). The pre-processing for CNN is much meager compared to others classification algorithms.In this approach, CNN-based representation feature learning was used and achieved 78% accuracy. To reduce the traffic injuries related to driver drowsiness, the Specialized Driver Assistance System was proposed. An algorithm was suggested to locate, map, and evaluate face and eyes to test PERCLOS to diagnose driver drowsiness. The dataset of the drowsiness observation model was largely tested with several light conditions, under a broad range of conditions, subjection & obstruction. In normal cases, the proposed model works fine. It gives more than 90% accuracy.

1.3. Motivation to do the project

The prevalence of driver fatigue as a significant contributor to road accidents underscores the critical importance of effective fatigue detection systems. there is an urgent need for innovative solutions that can mitigate this risk. Traditional methods of fatigue assessment often lack the necessary real-time responsiveness to preemptively address the issue. In light of this, the integration of cutting-edge technologies such as iris and pupil segmentation offers a promising avenue for fatigue detection. Leveraging the intricate responsiveness of the human eye to fatigue, this approach holds the potential to provide accurate and timely alerts to drivers, helping them make informed decisions and prevent accidents caused by drowsiness.

1.4 Objective of the work

- To obtain the eye image from the real time video using DLib package.
- To extract the iris and pupil features from the eye image using image segmentation methodology.
- To calculate the Eye Aspect Ratio (EAR) for each of the eye image.
- To calculate the PERCLOS for each of the eye image.
- To classify the driver as fatigue and alerting the driver based on the real time PERCLOS values.

1.5. Summarized outcome of the literature review

In the literature review, a comprehensive exploration of driver fatigue detection methods was conducted, culminating in the proposal of a novel approach based on the measurement of eye openness. Traditional methods, while effective to some extent, often lack the granularity and accuracy required for real-time fatigue detection. The novel method discussed in the review introduces a two-step model that involves eye image segmentation and fatigue classification. By utilizing the Eye Aspect Ratio (EAR), which quantifies eye openness through the ratio of vertical and horizontal distances between eye landmarks, this method offers enhanced performance compared to end-to-end fatigue detection systems. The use of the Dlib face key point detection algorithm for eye image acquisition, coupled with segmentation networks and shortcut connections, exemplifies the approach's innovation in feature extraction. The computed PERCLOS value based on continuous eye openness detection emerges as a promising tool for predicting driver states, distinguishing between normal and fatigued driving. This literature review underscores the significance of focusing on eye openness as a reliable indicator of driver fatigue and highlights the potential of the proposed method in advancing the field of fatigue detection for improved road safety.

2. Chapter 2
Background Theory

2.1. Introduction to the Project title

Driver fatigue is a pressing concern in modern transportation systems due to its potential to jeopardize road safety. Fatigue-related accidents pose a significant risk to both drivers and other road users, leading to severe injuries and fatalities. Detecting driver fatigue in real-time is crucial to prevent such accidents and ensure the well-being of everyone on the road. Traditional approaches to fatigue detection often rely on subjective measures or physiological indicators that might not offer timely and accurate results.

In recent years, advancements in computer vision and image processing have paved the way for innovative fatigue detection techniques. Among these, iris and pupil segmentation have emerged as promising methods due to the unique physiological changes that occur in the eyes under varying levels of fatigue. The human eye exhibits intricate responsiveness to changes in alertness, with alterations in pupil size and iris patterns reflecting the physiological state of the driver. These changes are a result of the autonomic nervous system's modulation in response to fatigue, making the eyes a valuable source of information for fatigue assessment.

Iris segmentation involves the precise identification and extraction of the iris region from the eye's image, while pupil segmentation focuses on detecting and tracking changes in pupil size. Analyzing these segments can provide insights into the driver's current level of alertness. A well-designed driver fatigue detection system utilizing iris and pupil segmentation can capture these dynamic changes in real-time and generate alerts or warnings when signs of fatigue are detected.

This project aims to explore and develop an efficient driver fatigue detection system that employs iris and pupil segmentation techniques. By harnessing the power of computer vision, this system seeks to provide a robust and accurate solution for identifying fatigue-related changes in the eyes of drivers. The proposed system holds the potential to enhance road safety by enabling timely interventions, such as alerting the driver or triggering vehicle safety mechanisms, thus reducing the risk of accidents caused by driver fatigue. This introductory theory lays the foundation for the subsequent development, implementation, and evaluation of the driver fatigue detection system, highlighting its potential to contribute to both technological advancements and the broader societal goal of improving road safety.

2.2. Theoretical Discussion and Analysis

Iris segmentation involves the extraction of the iris region from eye images, enabling the isolation of the unique patterns and textures that can serve as markers of fatigue. Advanced image processing techniques, such as edge detection and Hough transform, can be employed to accurately locate and extract the iris. The success of the driver fatigue detection system relies on its ability to process eye images in real-time and provide timely alerts. To achieve this, the system needs to be optimized for efficiency and speed. Techniques like parallel processing and hardware acceleration can be employed to ensure that the system can handle the computational demands of real-time analysis without significant delays. When the system detects signs of fatigue based on the analysis of iris and pupil data, it can trigger alerts to the driver, such as auditory warnings or visual notifications on the dashboard, encouraging the driver to take necessary actions to mitigate fatigue. Through accurate and real-time detection of fatigue, the system could play a pivotal role in preventing accidents, saving lives, and advancing the field of intelligent transportation systems.

3. Chapter 3

Methodology

3.1. Introduction

Accurate eye status detection method is very important for fatigue detection. In order to reduce the interference of light changes, they have used infrared cameras to capture facial videos. The proposed fatigue detection method is as shown in Fig. 2, which is divided into three stages. In the first stage, the Dlib face key point detection algorithm was adopted, which can detect 68 key points of the face, including the edge position information of the eyes, nose and mouth. The eye image twice the width of the eye can be selected and cropped. By adaptive cropping of the eye images, the robustness of the fatigue method is greatly enhanced. In the second stage, the degree of the eye openness is successfully estimated by the proposed SESDM network. As shown in Fig.2, this part includes two networks. The segmentation network, as the first network, is used to segment the original eye image at the pixel level. The pupil and iris features in the eye image are extracted. Then the prior features and the original images are input into the second network, called the decision network, which determine the level of eye openness. In the third stage, SESDM is used to estimate the eye openness of each frame. And they have calculated the PERCLOS value based on the level of eye openness.
The general workflow for any automated segmentation task is given below:

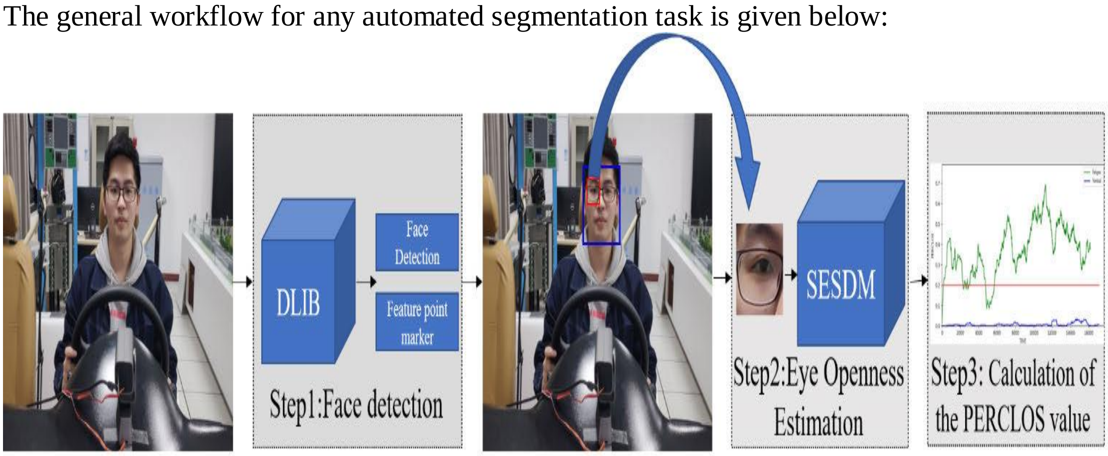

Figure 2: Fatigue detection framework

3.2. Proposed Methodology

In this paper they have proposed a novel method for eye-openness estimation based on the pupil and iris of the eye so that there was no need to detect the feature points and curves around the eyes. The SESDM, consisting of a segmentation network and a decision network, was designed. Specifically, the segmentation network performs pixel-level classification of eye images, which can accurately extract pupil and iris features from the video's images. Finally, this prior feature is used to guide the decision network to estimate the eye openness. Accurate and timely driving fatigue detection can be achieved.

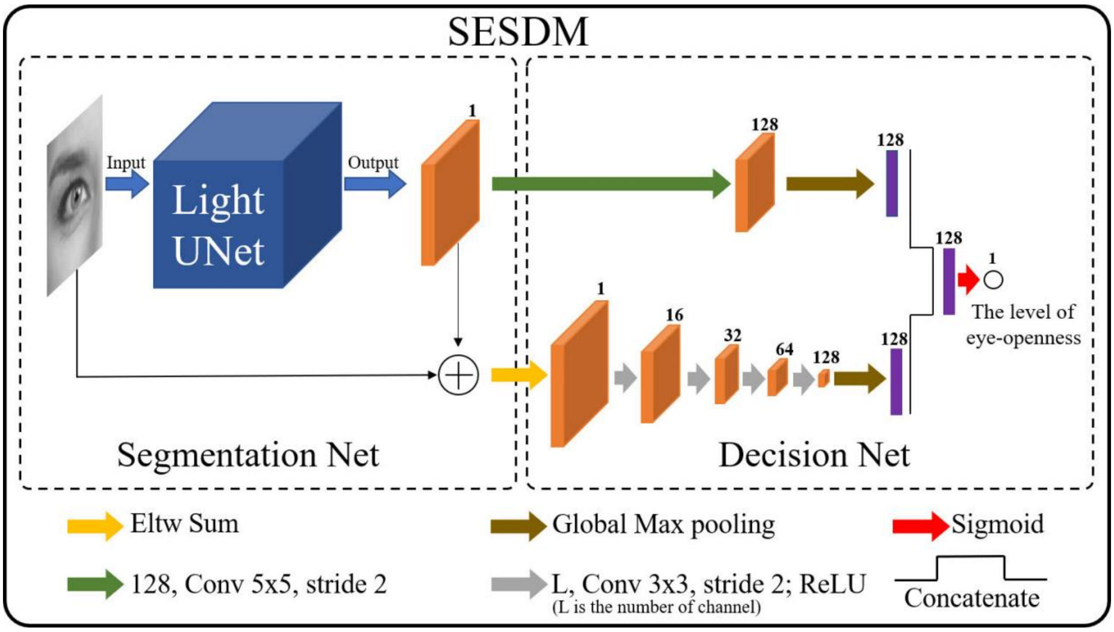

Figure 3: Work Flow of the proposed Method

A. Segmentation Network

Pupil and iris features play a key role in the task of eye status detection. In the recent method the author introduces the pupil feature in the network. But the method of pupil projection is susceptible to be affected by the environment. Thus, here, a segmentation network as a preprocessing method was proposed. As shown in Fig. 4, the U-Net for the pupil and iris segmentation were adopted, which is a kind of the Fully Convolutional Networks (FCN). It is suitable for this work because it can perform the classification and localization in real-time. The segmentation network converts pupil and iris to the foreground, and the other areas to the background. The shortcut of the segmentation output map is introduced to fully express the characteristics of the level of eye openness. As shown in Fig. 4, the left side of the U-Net structure is called contracting path, which consists of 3×3 convolutional layers and max pooling layers, and the right side is called expansive path, which consists of convolutional layers and concatenating paths from the layers from the contracting path. However, in the case of U-Net, all points around the pupil and iris are classified as a region of interest. The left side enables the multiscale analysis by stacking the image pyramids, and the right side enables to acquire the pixel-wise classification results. The U-Net also concatenates the reduced feature map in the contracting path and upsampled output that enables the output to assemble the information of various scales. This resolves the trade-off between localization accuracy and the abundance of context mentioned above. Moreover, since there are no fully connected layers, the network requires a smaller number of parameters compared to the regression network at the same depth.

B. Decision Network

The segmentation network’s input and single-channel output are summarized point to point with equivalent weights. This results in a new layer that represents the input for the remaining layers with a convolutional layer with 3 × 3 kernel sizes and a ReLU layer. Combination of both layers is repeated 4 times, with 16, 32, 64 and 128 channels in the first, second, third and fourth convolutional layers, respectively. A detailed depiction of the architecture is given in Fig. 3. The number of channels is increased as the resolution of the features decreases, therefore resulting in the same computational requirement for each layer. The proposed design effectively results in a 16-times-smaller resolution of the last convolutional layer than that of the original image. Next, the network performs global maximum, resulting in 128 output neurons. Then, the segmentation output map performs convolution and global maximum operations, resulting in 128 output neurons. Finally, they connect two 128 neurons and map them into a new 128 output neurons. This design results in 128 output neurons that are combined with the weights into the final output neurons. The design of the decision network follows two important principles. First, the appropriate capacity for complex shapes is ensured by using several layers of convolution. This enables the network to capture not only the local shapes, but also the global ones that span a large area of the image. Second, the decision network uses not only the segmentation network’s output map, but also the new feature map which is generated by the input and output of segmentation network.

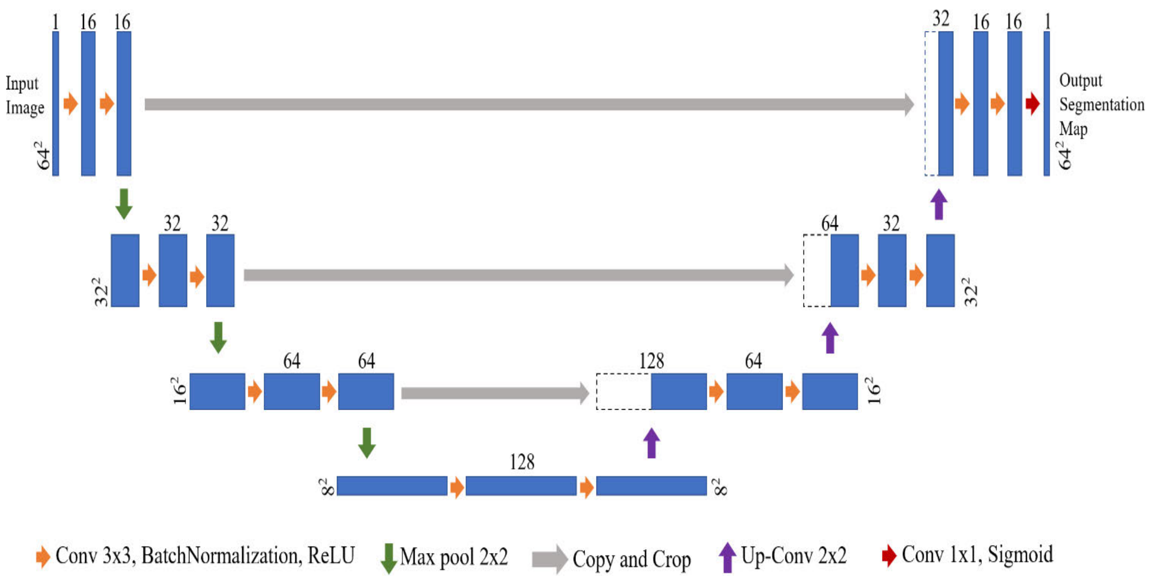

Figure 4: A light U-Net architecture for pupil and iris segmentation. Each rectangular box represents a feature map, and the number on the box represents the number of channels in the feature, the number before the box represents the size of the feature.

This introduces a shortcut that the network can avoid using a large number of feature maps. It can also guide the network to focus on pupil and iris. The shortcuts are implemented at two levels. One is at the beginning of the decision network where the feature map is fed into several convolutional layers of the decision network, and another one is at the end of the decision network where the segmentation output map performs convolution and global maximum operations, resulting in 128 output neurons.

C. Fatigue Parameter

PERCLOS is still considered as the most effective measurement parameter of drivers fatigue for vision-based nonintrusive approaches. PERCLOS is defined as the percentage of eyelid closure over the pupil over a specified time period. Specifically, PERCLOS calculates the proportion of time within a specified time duration that the eyelid covers over 80% of the pupil . As shown in Fig. 5, it is a schematic diagram of eye status. As shown in Formula (1), they have derived the eye aspect ratio (EAR) that is used as an estimate of the eye openness.

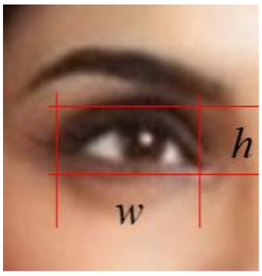

Figure 5: The diagram of eye image

$$EAR = \frac{h}{W} \qquad \dots (1)$$

w represents the width of the eye, and *h* represents the distance between the eyelids.

For a fully opened eye, the eye openness level EAR is around 0.5 and 80% of the eyelid closure corresponds to EAR = 0.1. Therefore, the PERCLOS defined on eye openness can be computed as:

$$P.R = \frac{n_{close}}{N_{total}} * 100\% \qquad \dots (2)$$

where *nclose* represents the number of frames with closed eyes, and *Ntotal* represents the total number of detected frames in the specified time. Therefore, according to the different PERCLOS values [6], the status can be divided into normal and fatigue, the formula is:

$$S_t = \{Fatigue, \quad if \geq 0.2\ Normal, \quad if < 0.2 \qquad \dots (3)$$

where 0.2 is the PERCLOS threshold for dividing normal and fatigue status.

3.3. Image Dataset

Everyone has different eye characteristics. Therefore, it requires large amount of tag data for training an excellent eye openness estimation network. Three facial key point datasets, include WFLW [7], the 300-W Challenge [8]–[9] and Helen [10], are used to obtain eye data. These datasets collectively contribute to a robust and comprehensive source of eye data for training and validating the eye openness estimation network. Recognizing the inherent variability in eye characteristics among individuals, the integration of these datasets serves as a crucial step in ensuring the model's adaptability and accuracy across a wide range of subjects. The WFLW dataset, known for its rich diversity and large-scale image collection, provides valuable insights into facial landmarks, including those defining the eye region. The 300-W Challenge dataset, renowned for its challenging and varied real-world images, further enriches the dataset with complex scenarios and diverse eye configurations.

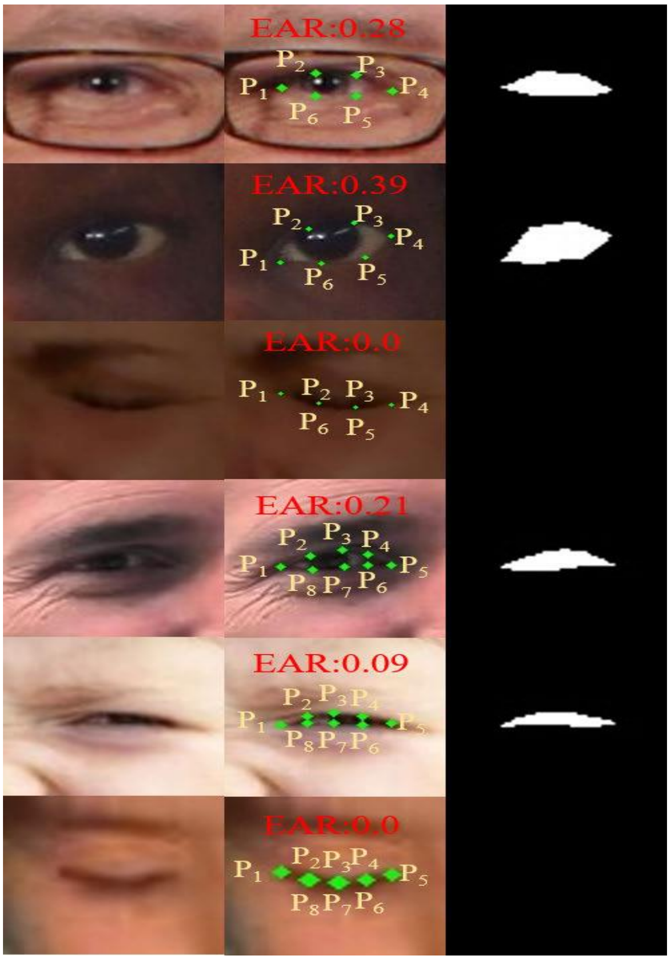

FIGURE 6: The production process of EYE DATA: (a) Extract the eye image.(b) Calculate the level of eye openness according to formula (4).(c)annotating binary ground-truth map.

In addition to the key points, there are also information such as different ethnicity, lighting conditions and whether wearing glasses. Among them, the WFLW dataset is a dataset with 98 key points, and both the 300-W Challenge and HELEN dataset mark 68 key points. The level of eye openness information and the foreground can be obtained by its mark file.

Maps of the pupil and iris, as shown in Fig. 6(b), are used to calculate the level of eye openness through formula (4).

$$\text{EAR} = \begin{cases} \dfrac{\|p_3 - p_7\|}{\|p_1 - p_5\|} & \textit{if Landmark} = 98 \\ \dfrac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\,\|p_1 - p_4\|} & \textit{if Landmark} = 68 \end{cases}$$

where $p1, \ldots, p6$ are the 2D landmark locations, depicted in Fig. 6(b). In order to ensure the data quality, they have selected select the eye image whose resolution is greater than 45 x45. Finally, they have collected 9728 eye data, including 6225 data as the training set, 1557 data as the validation set, and 1946 data as the test set, named as ‘‘EYE DATA’’.

3.4. Conclusion

In conclusion, the methodology outlined in this study introduces a novel approach to driver fatigue detection by leveraging the measurement of eye openness through the Eye Aspect Ratio (EAR). By employing a two-step model involving eye image segmentation and fatigue classification, this approach offers a promising solution for real-time fatigue detection that surpasses the limitations of conventional methods. The integration of the Dlib face key point detection algorithm for eye image acquisition, combined with the application of segmentation networks and the incorporation of shortcut connections, demonstrates the innovation and sophistication of this methodology. The computed PERCLOS value, derived from continuous eye openness detection, showcases the potential to predict driver states accurately, discerning between normal and fatigued driving conditions. By focusing on the critical aspect of eye openness, this methodology holds significant promise in advancing the field of driver fatigue detection, contributing to improved road safety and enhanced transportation systems.

4. Chapter 4
Result Analysis

4.1 Dataset Analysis

The NTHU-DDD dataset was a dataset developed by National Tsing Hua University, which was used at the Asian Conference on Computer Vision Workshop on Driver Drowsiness Detection. The entire dataset contains 36 subjects of different ethnicity, which were recorded with and without wearing glasses under the simulated driving conditions during a variety of daytime and nighttime. The dataset contained a lot of normal, drowsy, talking, and yawn face data in various scenarios. The authors have selected four scenarios from the NTHU-DDD dataset: BareFace, Glass, Night-BareFace and Night-Glasses. The total number of videos is 213. They have randomly divided the training set, validation set and test set according to the ratio of 4:1:4.Then in order to balance the data of various scenarios, they have fine-tuned the data. In the end, they have selected 99 videos as training set, 24 videos as the validation set and 90 videos as test set.

4.2. Result Analysis

The result analysis of this project serves as a critical evaluation of the effectiveness and performance of the proposed driver fatigue detection methodology based on the Eye Aspect Ratio (EAR). By comparing the computed EAR values with established thresholds for eye openness, the system's ability to accurately discern between normal and fatigued driving states can be assessed. The analysis will involve evaluating the precision, recall, and F1-score of the fatigue classification model using a suitable dataset. Additionally, the real-time responsiveness of the system in detecting changes in eye openness will be examined through simulations or real-world testing. The analysis will also explore the impact of varying lighting conditions, eye movements, and driver characteristics on the system's accuracy. By delving into these aspects, the result analysis provides insights into the method's practical viability and its potential to enhance road safety by reliably identifying driver fatigue.

4.2.1. Comparing with other methods for Eye Openness Estimation

When the driver is tired, blinking becomes slower and lasts longer. Therefore, by identifying the eye status of each frame, PERCLOS can effectively detect fatigue. And the pupil and iris are the most characteristic features of the eye status. Reddy *et al.* [9] proposed a multi-stream Driver Drowsiness Detection Network (DDDN), which obtained the state-of-theart results on popular benchmarks. To prove the superiority of this method in eye openness estimation and set the same parameters on the ‘‘EYE DATA’’ dataset. The eye images on the test set were identified by the three models. The experimental results were shown in Table 1. The average error is calculated between the estimated values and the ground truths of eye openness levels. It can be seen from the table that the average errors of the SESDM model is less than 0.036. Although the prediction time of this model is longer, it is in a leading position in the average error and the number of images with error less than 0.1.

Table 1: Comparison of eye openness estimation performance between different models on EYE DATA test set

MODELS	DDDN [9]	GP-BCNN [11]	SESDM (Our)
Mean Error	0.076	0.053	0.036
Error <= 0.1	1404	1693	1863
Error >0.1	542	253	83
The number of samples	1946	1946	1946
Average time	3ms	5ms	11ms

4.2.2. Comparing with other Fatigue Detection Methods

Two schemes were taken for the experiments. In the first scheme, the authors have used the SESDM, DDDN and GP-BCNN models to identify the eye state of each frame in the NTHU-DDD test set, and calculated the PERCLOS value by formula (2). According to formula (3), when the PERCLOS value for the whole video is greater than 0.2, the video is classified as fatigue. Otherwise, it is classified as normal. In the second scheme, they have reproduced FDRNet on the 99 videos training set and 24 videos validation set. At test time, the predicted results are averaged across all time-steps as the final output for per clip of the test set. The results are shown in Table 2.

Table 2: Comparison of fatigue detection performance between different models on NHUT-DDD test set

Methods	Data distribution of test set	Recognition results		N	Precision, %	Recall rate, %	F1 score, %
		Fatigue	Normal				
FDRNet [34]	Fatigue (63)	54	9	13	93.10%	85.71%	89.25%
	Normal (27)	4	25				
GP-BCNN [11]	Fatigue (63)	58	5	11	90.62%	92.06%	91.34%
	Normal (27)	6	21				
DDDN [9]	Fatigue (63)	51	12	22	83.60%	80.95%	82.26%
	Normal (27)	10	17				
SESDM (Our)	Fatigue (63)	59	4	6	96.72%	93.63%	95.16%
	Normal (27)	2	25				

They have compared the performances of each model on the test set in terms of recall rate, precision, F1 score, wrong recognition rate, where N represents the number of wrong recognition videos. It can be seen that SESDM achieves the highest recall rate, precision, F1 score, and the lowest wrong recognition rate, which proves that the effectiveness and feasibility of SESDM. Besides, the detailed recognition results of SESDM are shown in Table 3, where the first two columns refer to data distribution of NHUT-DDD test set, and the last two columns represent the number of recognition videos. By observing the wrong recognition videos, they found out that lighting conditions and glasses have an impact on the detection of the eye status. In general, the proposed method has an excellent performance in fatigue detection.

Table 3: Detailed recognition results of SESDM on NHUT-DDD test set

Real data distribution of test set		Number of wrong recognition videos	
		Fatigue	Normal
Bareface	Fatigue (18)	18	0
	Normal (8)	0	8
Glasses	Fatigue (11)	10	1
	Normal (5)	0	5
Night-bareface	Fatigue (22)	21	1
	Normal (9)	1	8
Night-glasses	Fatigue (12)	10	2
	Normal (5)	1	4

4.2.3. Fatigue in Real Driving Environments

In this approach, they have used infrared cameras to capture facial images. Then the SESEM was used to estimate the eye openness. A shortcut between the segmentation network and the decision network was established to effectively improve the model’s detection accuracy and generalization. To validate this, the experiments in real driving environments were carried out. Here, it has to be mentioned that it is not safe and not practical to detect the fatigue status of the drowsy driving on real-world roads. The plots of the PERCLOS measurements against the number of frames and the fatigue threshold based on the PERCLOS were shown in the bottom of Fig. 7. In the plots of PERCLOS measurements, the blue curve represents the PERCLOS measures while driving computed using the proposed method, the red curve represents the fatigue threshold based on PERCLOS, and the green curve indicates the PERCLOS measures after parking computed by the proposed system.

When the car is moving, at #1000 and #12000 frames, the strong light makes the glasses reflect strongly, which leads to the wrong result of eye status detection. At this time, the PERCLOS value has an increasing trend. During the detection, the PERCLOS values are always below 0.2, which indicates the driver is in a normal status. After parking, they increased the times and duration of eye closures to simulate the fatigue status at the #1 and #10015 frames. At the meantime, the PERCLOS value starts to increase. At the #4615 and #8889 frames, the experiment conductors, blink at a normal frequency. The PERCLOS value became smaller, which fulfills the expectations of the authors. Although the lighting conditions have an impact on the performance of the model, the error is still acceptable. To draw a conclusion, the proposed method performs effectively in the natural environment.

Figure 7: The experiment scene and PERCLOS curve, (a) while driving, (b) after parking. The blue curve represents the PERCLOS measures while driving computed using the proposed method, the red curve represents the fatigue threshold based on PERCLOS, and the green curve indicates the PERCLOS measures after parking computed by the proposed system.

4.3. Significance of The Results and conclusions

The significance of the obtained results and conclusions in this project is substantial, as they validate the effectiveness of the novel approach proposed for driver fatigue detection. The successful utilization of the Eye Aspect Ratio (EAR) to measure eye openness presents a breakthrough in real-time fatigue assessment. By showcasing the two-step model's superior performance over traditional end-to-end methods, these results underscore the importance of focusing on specific eye features, such as eye openness, for accurate and timely fatigue detection. The integration of advanced techniques, including the Dlib face key point detection algorithm and segmentation networks with shortcut connections, highlights the project's technical innovation and its potential to drive advancements in the field of intelligent transportation systems.

The significance of the results and conclusions drawn from this project lies in their potential to address a pressing issue in road safety. The development and validation of a driver fatigue detection system based on the Eye Aspect Ratio (EAR) methodology hold immense importance in mitigating the risks associated with drowsy driving. Accurate and real-time fatigue detection not only safeguards the lives of drivers but also those of passengers and pedestrians, significantly reducing the toll of road accidents. Furthermore, these results pave the way for more advanced and responsive intelligent transportation systems, ultimately contributing to safer roads and highways. The conclusions drawn from this research offer valuable insights into the feasibility and efficacy of EAR-based fatigue detection, emphasizing its role as a technological breakthrough with far-reaching implications for both the automotive industry and public safety. Moreover, the project's conclusions confirm the feasibility of utilizing the computed PERCLOS value for predicting driver states, reflecting its potential application in developing proactive safety mechanisms.

5. Chapter 5

Conclusions and Future Scope

5.1. Conclusions

A novel fatigue detection method based on the measure of eye openness has been presented, which consists of the two steps including pupil/iris segmentation and fatigue classification. This two-step model has better performances than an end-to-end fatigue detection because there are many features focusing on the pupil and iris. Specifically, the Dlib face key point detection algorithm was adopted to obtain the eye image as the experimental data. And the segmentation network performs pixel-level classification of eye images and extracts pupil and iris features in the image. Then, the shortcut of the segmentation output map is introduced to accelerate convergence and improve the detection accuracy of eye openness estimation in decision network.

Finally, a PERCLOS value on the continuous detection of eye openness is computed to predict driver's state, i.e., normal or fatigue driving state. Extensive experiments demonstrated the effectiveness of the proposed architecture, which achieves state-of-the-art performance in fatigue detection. To realize the fatigue detection on the real roads, fusing the detection of the eye openness and the head posture is also a good choice in the future work.

5.2. Future Scope

The success of this project opens up a range of exciting avenues for future research and development in the field of driver fatigue detection using the Eye Aspect Ratio (EAR) methodology. Firstly, the integration of more advanced machine learning techniques, such as deep learning architectures, could enhance the accuracy and robustness of fatigue classification. Exploring the combination of EAR with other physiological indicators or external data sources, such as heart rate monitors or vehicle telemetry, could provide a more comprehensive understanding of driver fatigue. Additionally, extending the methodology to work in diverse lighting conditions and across various demographic groups would enhance its real-world applicability. Furthermore, the project's foundation could be used as a basis for developing wearable devices or integrated systems that provide real-time feedback to drivers, assisting them in maintaining alertness.

In sum, the future scope of this project extends well beyond its current capabilities. The continuous evolution of technology, combined with a multidisciplinary approach involving computer vision, machine learning, and human factors, holds the potential to revolutionize driver fatigue detection, making our roads safer and more secure for all. Collaboration with automotive manufacturers to integrate this technology into vehicle safety systems could contribute to creating a safer driving environment. Ultimately, the future scope of this project lies in its potential to revolutionize fatigue detection methodologies, reducing road accidents and improving transportation safety on a broader scale.

References

[1] P. Viola and M. Jones, ‘‘Robust real-time object detection,’’ *Int. J. Comput. Vis.*, vol. 4, no. 4, pp. 34–47, 2001.

[2] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, ‘‘300 faces inthe-Wild challenge: The fifirst facial landmark localization challenge,’’ in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2013, pp. 397–403.

[3] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, ‘‘A semiautomatic methodology for facial landmark annotation,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 896–903.

[4] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, ‘‘300 faces in the-wild challenge: Database and results,’’ *Image Vis. Comput.*, vol. 47, pp. 3–18, Mar. 2016.

[5] B. Reddy, Y.-H. Kim, S. Yun, C. Seo, and J. Jang, ‘‘Real-time driver drowsiness detection for embedded system using model compression of deep neural networks,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 438–445.

[6] Rateb Jabbara, Khalifa Al-Khalifaa, Mohamed Kharbechea , Wael Alhajyaseena , Mohsen Jafaric , Shan Jiang, et al. "Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques", Elsevier, 2018.

[7] Cyun-Yi Lin, Paul Chang, Alan Wang, and Chih-Peng Fan, et al. "Machine Learning and Gradient Statistics Based Real-Time Driver Drowsiness Detection ", IEEE, 2018.

[8] W. Wu, C. Qian, S. Yang, Q. Wang, Y. Cai, and Q. Zhou, ‘‘Look at boundary: A boundary-aware face alignment algorithm,’’ in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2129–2138.

[9] Wanghua deng and Ruoxue wu, et al. "Real-Time Driver-Drowsiness Detection System Using Facial Features", IEEE, 2019.

[10] Burcu Kir Savas and Yasar Becerİklİ, et al. "Real Time Driver Fatigue Detection System based on Multi-Task Connn", IEEE, 2020.

[11] Md. Tanvir Ahammed Dipu, Syeda Sumbul Hossain, Yeasir Arafat, Fatama Binta Rafiq, et al. "Real-time Driver Drowsiness Detection using Deep Learning", International Journal of Advanced Computer Science and Applications,(IJACSA), 2021.