https://doi.org/10.2352/ISSN.2470-1173.2021.8.IMAWM-233

Real-time Detection of Early Drowsiness Using Convolution Neural Networks

Chinh Tran, and Nader Namazi
Department of Electrical Engineering and Computer Science, The Catholic University of America, Washington, DC 20064

Abstract

Drowsiness driving is one of the major reasons causing deadly traffic accidents in the United States of America. This paper intends to propose a system to detect different levels of drowsiness, which can help drivers to have enough time to handle sleepiness. Furthermore, we use distinct sound alarms to warn the user to prevent early accidents. The basis of the proposed approach is to consider symptoms of drowsiness, including the amount of eye closure, yawning, eye blinking, and head position to classify the level of drowsiness. We design a method to extract eye and mouth features from 68 key points of facial landmark. These features will help the system to detect the level of drowsiness in real-time video stream based on different symptoms. The experiential results show that the average accuracy of the system that has the capability to detect drowsiness intensity scale in different light conditions is approximately 96.6%.

Introduction

Drowsiness driving is one of the major problems worldwide and especially in the United States of America. According to National Highway Traffic Safety Administration (NHTSA) statistics, around 90,000 crashes caused from drowsiness driving between 2015-2017, while the reported deaths approached 4000 people from 2013-2017 [1]. There are several reasons making people sleepy while driving; one of the studies shows that driving for a long period of time makes the driver lose their self-judgment and concentration [1]. Sleepiness will affect driver's ability to observe surrounding things to drive safely. However, Drivers will not stop driving even if they fall asleep, most of them will say "I will be fine, I can continue driving". According to National Sleep Foundation (NSF), there are some previous signs of drowsiness that can alert a driver to stop and rest, such as, frequent blinking, yawning repeatedly, eye closure continuously, mouth opening, and/or keeping his/her head up [2].

Previous work in this field mostly focused on detecting one or more of these signs combined. Some researchers focus on detecting yawning [3][4] while others concentrate on eye detection [5][6]. Publications that considered yawning detection mostly measured the width and length of the open mouth. Wang in [3] detects a ratio between two perpendicular lines running through upper and lower lip boundaries. Moreover, he detects yawning when this ratio is larger than a threshold of 0.5 in more than 20 consecutive frames in a 30 frames per second video. Lu in [4] detects yawning by considering the geometric feature changes in two consecutive frames using the vertical distance between the nose center and the chin. However, these algorithms cannot detect whether people are opening their mouth for a normal activity (speaking, smiling, singing, etc.) or for yawning. On the other hand, eye detection is another important sign for detecting drowsiness. Some papers proposed for counting blinking eyes by using Support Vector Machine (SVM) and Eye Aspect Ratio (EAR) [5][6]. However, the accuracy of these methods suffers in different illumination conditions. Lighting conditions may change dramatically while driving, which can affect the accuracy of eye detection. In addition, the use of glasses should be considered while driving. Awasekar in [26] combined both features for drowsiness detection, but they used a fixed threshold for eye detection which might vary from one person to another causing false detections.

In this work, drivers with/without glasses and continuously changing light conditions are considered. Moreover, a combination of both eye and mouth movements are adopted to detect the level of drowsiness for distinct individuals. Media Research Lab (MRL) eye dataset that included opened and closed eyes with/without glasses, good/bad lighting conditions, and none/small/big reflection for both left and right eyes is selected for this project [7].

This research focuses on analyzing the level of sleepiness by proposing a novel Convolutional Neural Networks (CNN) to detect drowsiness. This network is used to combine features of the mouth and eye regions with the least possible complexity. According to NSF symptoms (Table 1), three different kinds of drowsiness levels are specified [2]. As shown in Table 1, an opening eye without yawning can be described as a normal driver with no drowsiness detected. On the other hand, frequent blinking and yawning can indicate a less drowsy driver. The most severe drowsiness level is a result of closed eyes over 1.5 seconds.

Table 1: Level of drowsiness

Symptoms	Output
Eye open, no yawning	No drowsiness
Frequent blinking, yawning	Less drowsiness
Eye closed over 1.5 seconds	Drowsiness

An overview of the related study for drowsiness detection, input selection and drowsiness methods are explained in the related work section. Then, system overview will provide the proposed system workflow for classifying eye and mouth state and detecting level of drowsiness in real-time video stream. Next, experimental results are illustrated followed by the last section on conclusions.

Related works

Several researchers have studied drowsiness detection in the past few years. An overview of their datasets and drowsiness methods are mentioned in this section. The pros and cons of selecting a dataset that is suitable to train our proposed CNN model is also included.

Dataset Selection

Many datasets have been applied in various papers, however, most of datasets are neither public nor realistic. Therefore, each work uses a different dataset to solve drowsiness detection. Unfortunately, the lack of standardized dataset results in an infeasible comparison between different algorithms. Several methods achieve high validation accuracy of 91.6% [8] and 95.8% [9] but with using private datasets, others use public datasets, but they are unrealistic and lack the number of sufficient training samples [12, 13], such as, ULG Multimodality Drowsiness (DROZY) database [12]. Another obstacle while using such a dataset is related to image processing tasks, including cropping, image resizing and accuracy recognition. These tasks are affected by the location of the sensor that is attached to the subject ‘s face. On the contrary, some datasets [10,11], which were only applied for detecting human’s emotion through facial expression, does not serve the drowsiness detection problem. Another dataset is National Tsing University (NTHU) driver drowsiness dataset [13] contains both RGB and IR video for 36 participants in many driving scenarios. However, these subjects pretend to be drowsy and their environment condition is unrealistic and insufficient samples which affect drowsiness detection in real situations.

The Closed Eyes in the Wild (CEW) dataset [14] has a huge number of images representing the eyes of 2423 subjects that is randomly selected from the internet. Nevertheless, these images are created from an impractical environment conditions as well as IR reflections are neglected. Zhejiang University (ZJU) dataset [15] is deployed in [9], which is collected from 20 individuals under natural light conditions. However, this dataset contains only indoor image without IR reflections.

In comparison to another datasets, MRL eye dataset is based on the quality of the images that corresponds to the conditions that occur in the real environment. These conditions strongly rely on the ambient light and/or the variations of the distance between the camera and the driver [7]. This dataset has different kinds of reflections (no reflections, weak reflections, and strong reflections) and light conditions (good or bad) as illustrated in Figure 1. These samples are collected from 37 distinct subjects with/without glasses for both left and right eye. Moreover, MRL eye dataset is based on manually cropped eye region images that is perfectly appropriate to be represented as an input for our proposed CNN model.

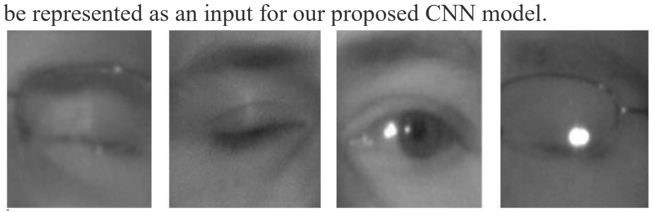

Figure 1: Example of eye reflections with/ without glasses dataset [7]

Drowsiness Detection Methods

Various techniques have been reported in the literature to detect drowsiness in the past few decades [6,16,17,18]. This section will review these techniques used to classify the level of drowsiness. Eye feature is an important feature to detect drowsiness driving. Martin and Carvalho [19] show the relationship between blinking and fatigue by using Electroencephalographic (EEG), Electrooculographic (EOG), eye tracking and video camera systems. Friedrichs and Yang [20] extract some features for drowsiness detection. These features include blink duration, eye closure, blinking rate, mean eye opening, microsleep and head nodding frequency for subjective self-estimation by Karolinska Sleepiness Scale (KSS). However, all these methods are limited because they require additional hardware that is impractical.

Recent researches focus on applying machine learning and deep neural network to detect drowsiness driving. First of all, SVM are used to learn and classify different states of drowsiness from labeled data with high accuracy as in [16,17]. In [16], the author proposed a database with 170 feature state EAR samples with 3 different labels (open-0, close-2, blink-1). SVM use EAR from latest set of 15 consecutive frames to predict output of every 7 new frames received from the video feed with 0.28 seconds measurement. However, finding the best representation for the mapping to separate each region for large data is complex. It requires a set of parameters in SVM formula to provide proper values for training model. Furthermore, EAR cannot be used as a general signature since each person has a distinct EAR. Similarly, Park et al [17] proposed a new architecture using three deep neural networks and applied an SVM to the combined features of those three networks to classify four classes of alert, yawning, nodding and drowsy with blinking for each frame. However, the work of Park only focuses on improving drowsiness detection accuracy and not considering performance in terms of speed for real time applications. Novie et al [6] use EAR in [18] to find the threshold for an individual driver, however, because each driver has different EAR, this approach predicts poor results. Another disadvantage is that normal activities such as smiling, talking will affect EAR results and bring false alarms for drivers. The following section will concentrate on reviewing the entire system.

System Overview

This section provides a CNN model for the classification of eye state and the detection of drowsiness level in a real-time video based on eye and mouth symptoms. Figure 2 illustrates system workflow and the procedure of detecting drowsiness. The first stage in the system workflow represents the overview of the offline learning process. This stage briefly describes the steps that are followed for the CNN training process to classify opened and closed eye. After the offline training process, the eye and mouth states in real-time video can be detected via the online operating process. As a result, the level of drowsiness is predicted based on symptoms from these states.

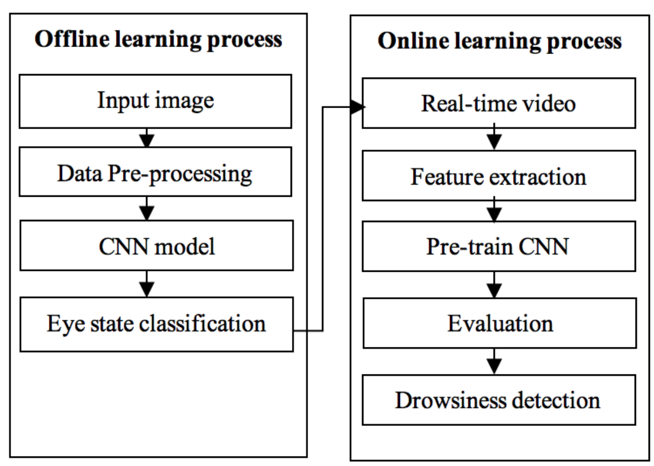

Figure 2: System workflow

Offline Learning Process

This section is concerned with the description of the process for classifying opened and closed eyes by using the proposed CNN offline learning model, as shown in Figure 2. With CNN, the system can detect eye closure, blinking frequency that describes symptoms of drowsiness. A more comprehensive explanation of the learning CNN process is provided next.

CNN Model

A proposed CNN architecture that includes convolutional layers, pooling layers and fully connected layers is illustrated in Figure 3. The total number of layers used to classify eye state are ten layers. There are three convolutional and max pooling layers, one drop-out and fully connected layer. The convolutional layer will extract useful features from the input image. Each convolutional layer is connected to an RELU activation to increase the non-linear features in each image. Then, a max pooling layer is used, not only to preserve the main features, but also to reduce the size of the images. This helps CNN model reduce the amount of unimportant information while classifying the eye state image. Next, a dropout layer is deployed, this layer provides a way of combining many different neural network layers efficiently and prevents from overfitting. After That, a fully connected layer transforms the entire feature matrix received from the previous layers to a single column vector that represented the important features. After flattening, the single column will be processed by the RELU activation. This procedure increases the features non-linearity to enhance the soft-max decision. Finally, a single neuron with sigmoid function is added for our binary classification (0,1 closed and open eye respectively) [21].

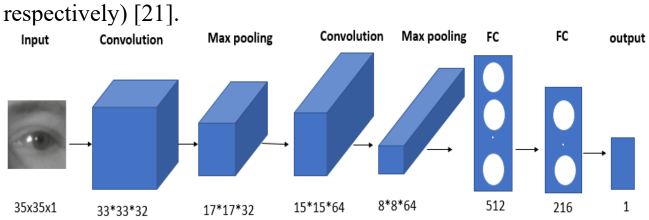

Figure 3: Architecture model of proposed CNN

Online Learning Process

After the offline training process, we describe the online process of detecting the level of drowsiness as illustrated in Figure 4. First, we need to consider light condition adjustments. Then, features are extracted from the real time video input. These features are fed to the pre-train CNN model. In addition, Mouth Aspect Ratio (MAR) is used to evaluate symptoms of drowsiness from these features. As a result, a classification of the level of drowsiness is detected and an audio warning is announced to alert the drivers.

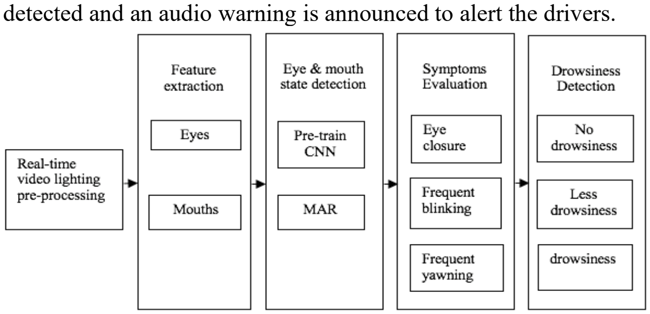

Figure 4: An overview of the online learning process

Real-Time Video Lighting Pre-Processing

During driving drowsiness detection in real time, the quality of the image is an important factor that needs to be considered. Different weather conditions including rain, fog, snow, or darkness affect the light conditions, which influence the system's drowsiness detection. For this purpose, each image frame is tested by using perceived brightness to decide if light enhancement is necessary. The perceived brightness Y is calculated by finding the average driver's facial pixels of RGB color spaces as shown below:

$$\left\{ \begin{array}{c} R,G,B = \frac{\sum_n^m H}{m-n+1} \\ Y = 0.2126R + 0.7152G + 0.0722B \text{ [22]} \end{array} \right\} \quad (1)$$

Where H is the value of RGB color spaces, whereas n and m represent the first and last pixels of the driver's facial frame, respectively, and R, G, and B express the three different channels.

The range of brightness can be estimated by calculating Y for many images. The estimated value of Y will be an indication of an image to be dark, bright, or extremely bright when $0 < Y < 120$, $120 < Y < 200$, and $Y > 200$, respectively. Only two scenarios are considered for light adjustment: namely, dark, and extremely bright scenarios. In these cases, we use histogram equalization [23]. As shown in Figure 5, the equalization increases the brightness of the dark image frames and decreases the level of brightness in extremely bright image frames.

Figure 5: (a) Original dark image, (b) Dark image after applying histogram equalization, (c) Original extremely bright image, (d) Extremely bright image after applying histogram equalization

Feature Extraction

Facial landmarks are used to localize and label salient regions of the face such as: eyes, eyebrows, nose, mouth, and jawline. The location of 68 coordinates (x, y) that maps to facial structure is shown in Figure 6a. This paper considers only eye and mouth regions to detect driving drowsiness. To proceed, ensemble regression trees for facial landmark detection is used to detect the height and width eyes as reported in [24]. To extract the area of the eye, 12 points are considered to find the eye bounding box. To do this, first, the face must be detected to extract the face bounding box. Face cascade is chosen because it is faster than other methods in terms of detection, tracking and efficiency [25]. Figure 6b illustrates the approach to find the left eye bounding box ABCD. Here, p1, p2, p3, p4, p5 and p6 are 6 coordinates (x, y) eye landmark ratio. Following the same method as eye extraction, we use point 49 to point 62 as the mouth location. This is illustrated in Figure 6c.

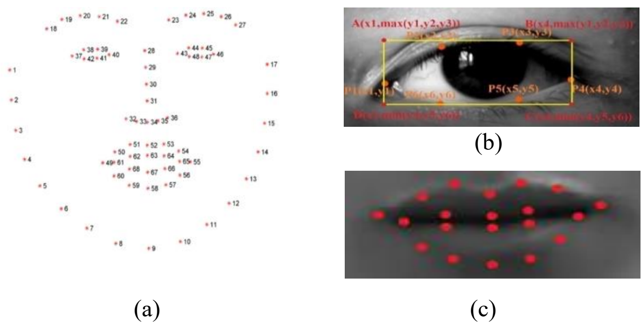

Figure 6: (a) Visualizing 68 facial coordinate points from the iBUG 300-W dataset [24], (b) Eye extraction, (c) Mouth extraction

Eye and Mouth State Detection

After extracting both eye regions from real-time video, the left eye region (ABCD bounding box) is resized. This procedure is applied to fit the image input layer of the CNN model for training that is predefined to 35x35 pixels. These coordinators are calculated as follows:

$$\text{Coordinate of Calibrated A} = \left(\frac{x1+x4-35}{2}, \max\,(y1, y2, y3) - \frac{x1+x4-35}{2}\right) \quad (2)$$

$$\text{Coordinate of Calibrated B} = \left(\frac{x1+x4+35}{2}, \max\,(y1, y2, y3) - \frac{x1+x4+35}{2}\right) \quad (3)$$

Similarly, coordinate of calibrated D and C are calculated comparable to A and B respectively, with min (y4, y5, y6) and min (y1, y2, y3). As shown in Figure 7, 6 points out of the 20 points are focused on selecting that were described in the previous paragraph. From these 6 points, the MAR (Mouth Aspect Ratio) equation is calculated as [26]:

$$MAR = \frac{|BF| + |CE|}{2 * |AD|} \quad (4)$$

Using the value of MAR, we can set the threshold to classify open/close mouth. According to our measurements of twelve volunteers, the average yawning threshold is set to 0.79.

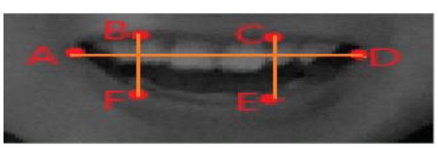

Figure 7: Architecture model of proposed CNN

One of the technical challenges for detecting MAR is its ability to distinguish between normal (smiling, talking, singing, laughing) and yawning mouth expression. To handle this defect, yawning can be detected when the driver's mouth opens for a certain amount of time. A graphical representation shows different mouth activity per frames, as illustrated in Figure 8. According to [27], a driver is considered yawning when the mouth is open for 7 seconds. Therefore, a driver is assumed to be yawning when his/her mouth has been opened for consecutive frames higher than the threshold of 7 seconds.

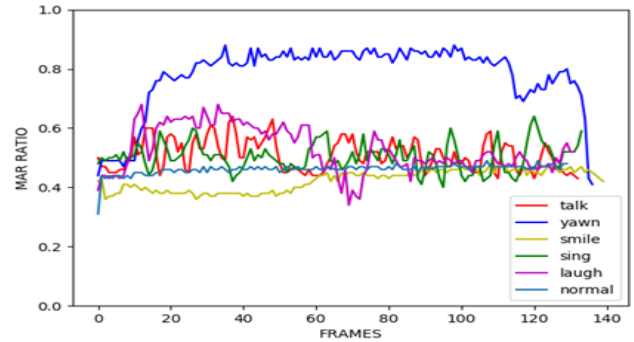

Figure 8: MAR in different mouth state

Drowsiness Detection and Alert System

There are three noticeable symptoms to detect drowsiness, as presented in Table 1. Based on the different levels of symptoms, drowsiness is classified. The formulas for detecting these symptoms are explained below.

Eye closure refers to the number of frames where closed eyes are detected in a certain time frame. In [28], a driver is considered drowsy when the number of consecutive eye-closed frames exceeds 45 (1.5 seconds). The eye blinking is detected when the state of the eye changes from opened to closed. In the literature, it is specified that a normal person would blink around ten times per minute [28]. If the eye blinking frequency is less, the driver will be considered drowsy. The equation of drowsiness frequent blinking eye is given by:

$$F = \frac{E}{M} \quad (5)$$

where E is the number of blinking eyes and M is the number of frames in a minute. Yawning detection occurs when an open mouth is notable. The equation of frequent yawning is computed by:

$$Y = \frac{n_y}{N} \quad (6)$$

where n_y is the number of frames with yawning detection and N is the total number of frames per minute.

Distinct alarms are generated to warn the user to wake up for different drowsiness level. The following algorithm summarizes the proposed procedure implemented for drowsiness detection.

Algorithm 1: Drowsiness Detection Algorithm

Input: real-time video
Output: Level of drowsiness detection
Extract frames from video
Adjust lighting conditions by using histogram equalization
Detect eye and mouth location
Identify eye status
Identify mouth status
Calculate C; number of consecutive eye closed
Calculate N; frequent blinking in 1 minute
Calculate M; frequent yawning in 1 minute
If C> 45 and not yawning, then
    The driver is drowsy, an alarm wakes the driver up
End
If F< 10 or Y>3 then
    The driver is less drowsy, alert messages with advice
Else
    The driver is awake
End

Experimental Results

This section includes a comparison for the offline training process between the proposed model and other models. The online operating process successfully validates the drowsiness level accuracy in real-time video stream.

Offline Learning Results

The training data has 1150 opened right eye and 1120 closed right eye image samples. The testing data contains 350 opened right eye and 280 closed right eye image samples. Table 2 lists the proposed CNN and other popular models that were used for offline learning. It shows that the proposed CNN model achieves the highest validation accuracy in a short training time with 10 epochs and 32 batch size.

Table 2: Accuracy comparison of the proposed CNN model to another models

Model	Validation Accuracy
Proposed CNN	**98.66%** (10 minutes)
VGG16 [30]	97.05% (22 minutes)
Alexnet [31]	97.15% (19 minutes)
Zhang et al [9]	97.85% (9.5 minutes)

Online Learning Results

Dell XPS 15 webcam with 16 GB memory is used to validate the accuracy of the model in a real-time situation. In addition, Python 3.6, OpenCV 4.0 [32], Keras 2.3 are used to build the software for validating the accuracy of the drowsiness detection. To test the accuracy of the system in different scenarios, we measure the precision of each symptom in various light conditions. Dlib library [24] is used for the facial landmark. Pysound Library [29] is used to generate distinct sound alarms to warn the user to wake up for different drowsiness level. The algorithm was tested on twelve different users, the results are shown in Table 3.

Table 3: Accuracy comparison of different symptoms for different light conditions

Symptom	Light Condition	Val. Acc.
Blinking alone without yawning	Darkness	96.2 %
Yawning alone without blinking	Darkness	97.0 %
Yawning and Blinking	Darkness	96.4 %
Blinking alone without yawning	Brightness	98.5 %
Yawning alone without blinking	Brightness	97.7 %
Yawning and Blinking	Brightness	97.5 %
Blinking alone without yawning	Extreme Brightness	95.3 %
Yawning alone without blinking	Extreme Brightness	96.0 %
Yawning and Blinking	Extreme Brightness	95.5 %

It is shown from Table 3 that our system performs more efficiently in bright light conditions. Table 3 also indicates that the system accuracy reduces in either dark or extreme bright light conditions. In conclusion, the average accuracy of the system that has the capability to detect severe level of drowsiness is approximately 96.6%.

Furthermore, according to distance measurements taken for twelve different drivers, our system performs efficiently when the distance is between 15.8" and 31.5". For this range, the average accuracy of the system achieves approximately 96.2%. Table 4 summarizes the accuracy of the system in different distances within this range.

Table 4: Different Distances for Bright Light Condition

Symptom	Validation Accuracy		
	15.8” < d < 20”	20” < d < 25”	25” < d < 31.5”
Blinking alone without yawning	97.2%	98.1%	96.7%
Yawning alone without blinking	96.2%	97%	95.8%
Yawning and Blinking	94.5%	95.6%	94.3%

Conclusions

This work has been concerned with the development and implementation of an algorithm to detect and alarm drowsy drivers. A proposed CNN model achieves higher accuracy than another models in offline learning process by selecting right eye from MRL eye dataset. We have extracted the eye and mouth features by using facial landmark. Based on eye and mouth states, we have designed a new evaluation method for defining symptoms of drowsiness. A system is proposed for detecting level of drowsiness hinged on these symptoms. Experimental results show that the novel system has high and stable performance in different light conditions. Furthermore, the system can operate with high speed to detect status of eye and mouth promptly in real-time video. To the best of our knowledge, our reliable real-time system can help the drivers avoid the drowsy-driving crashes. In future work, head position can be integrated as a factor to increase the robustness of drowsiness detection.

References

[1] Taylor, C.,“ Drowsy Driving Statistics”, https://www.thezebra.com/research/drowsy-driving-statistics/, 2020

[2] NSF, "Warning Signs”, https://drowsydriving.org/about/warning-signs/ ,2020.

[3] Wang, T. and Shi, P., “Yawning Detection for Determining Driver Drowsiness”, Proc. IEEE International Workshop, 2005.

[4] Lu, Y. and Wang, Z., “Detecting Driver Yawning in Successive Images”, 1st International Conference on Bioinformatics and Biomedical Engineering, 2007.

[5] Anizy, G.J., Nordin, M.J., and Razooq, M.M, “Automatic Driver Drowsiness Detection Using Haar Algorithm and Support Vector Machine Techniques”. In: Asian Journal Appl. Sci. Papers, vol.8, no.2, pp.149-157, 2015.

[6] Theresia, N., Pasaribu, Br., Ratnadewi, Prijono A., Adhie R.P., and Hermawan J.F, “Drowsiness Detection System Design based on Individual Driver”, Proc. AIP Conference 2097 ,2019.

[7] Fusek, R., “Pupil Localization Using Geodesis Distance”, ISVC 2018 Advance in Visual Computing, pp.433-444,2018.

[8] Reddy, B., Kim, Y., Yun, S., Seo, C. and Jang, J., “Real-time Drowsiness Detection for Embedded System Using Model Compression of Deep Neural Networks,” IEEE Conference on Computer Vision and Pattern Recognition Workshops ,2017.

[9] Zhang, F., Su, J., Geng, L. and Xiao, Z., “Driver Fatigue Detection based on Eye State Recognition”, IEEE Conference on Machine Vision on Information Technology ,2017.

[10] Yan, W., Wu, Q., Liu, Y., Wang, S., and Fu, X., “CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces”, 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition ,2013.

[11] Yan, W., Li, X., Wang, S., Zhao, G., Liu, Y., Chen, Y., and Fu, X., “CASME II: An improved spontaneous micro expression database and the baseline evaluation”, PLoS ONE ,2014.

[12] Massoz, Q., Langohr, T., Francois, C. and Verly, J., “The ULg Multimodality Drowsiness Database (called DROZY) and Examples of Use”, IEEE Winter Conference on Application of Computer Vision ,2016.

[13] Weng, C., Lai, Y. and Lai, S., “Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network”, In Asia Conference on Computer Vision 10118, pp.117-133 ,2017.

[14] Huang, G.B., Ramesh, M., Berg, T. and Miller, E., “Labeled Face in the Wild: A Database for Studying Face Recognition in Unconstrained Environments” ,2008.

[15] Pan, G., Sun, L., Wu, Z. and Lao, S., “Eyeblink-based anti-spoofing in Face Recognition from a Generic Web Camera”, IEEE 11th International Conference on Computer Vision ,2007.

[16] Caio, B., Marcio, C., Joao, S., Lucas, N., July, M., Isis., L. and Enrique, D., “Real-time SVM Classification for Drowsiness Detection Using Eye Aspect Ratio”, Proceeding of Probabilistic Safety Assessment and Management PSAM 14 ,2018.

[17] Park, S., Pan, F., Kang, S. and Yoo, C., “Driver Drowsiness Detection System based on Feature Representation Learning Using Various Deep Networks”, In Asia Conference on Computer Vision 10118, pp.154-164 ,2016.

[18] Soukupova, T. and Cech, J., “Real-Time Eye Blink Detection Using Facial Landmarks”, 21st Computer Vision Winter Workshop ,2016.

[19] Martins, R. and Carvalho, J., “Eye blinking as an Indicator of Fatigue and Mental Load – a Systematic review”, Occupational Safety and Hygiene III ,2015.

[20] Friedrichs, F. and Yang, B., “Drowsiness Monitoring by Steering and Lane Data based Features under Real Driving Conditions”, Signal Processing Conference ,2010.

[21] Turian, J. and Bergstra, J. and Bengio, Y., “Quadratic Feature and Deep Architectures for Chunking”, Proceeding of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp.245-248 ,2009.

[22] Anderson, M., Chandrasekar, S., Motta, R. and Stokes, M., “A Standard Default Color Space for the Internet: sRGB”, Technical Report International Color Consortium,1996.

[23] Wanghua, D. and Ruoxuw, W., “Real-Time Driver-Drowsiness Detection System Using Facial Features”, IEEE Access (7), pp.118727-118736 ,2019.

[24] Sagonas, C., Antinakos, E., Tzimiropoulos, G., Zafeiriou, S. and Pantic, M, “300 Faces In-The-Wild Challenge: Database and Results”, Image and Vision Computing (47), pp.3-18,2016.

[25] Mael, F.,“ A guide to Face Detection in Python” https://towardsdatascience.com/a-guide-to-face-detection-in-python-3eab0f6b9fc1, 2020.

[26] Awasekar, P., Ravi, M., Doke, S. and Shaikh, Z., “Driver Fatigue Detection and Alert System Using Non-Intrusive Eye and Yawn Detection”, International Journal of Computer Applications 180(44), pp.88-87 ,2018.

[27] Eguibar, J., Uribe, C., Cortes, C., Bautisa, A. and Gallup, A., “Yawning Reduces Facial Temperature in the High Yawning Subline of Sprague-Dawley Rats”, BMC Neuroscience 18(1), pp.3,2017.

[28] Theresia, N., Pasaribu, Br., Prijono, A., Ratnadewi, R., Adhie, R. and Felix, J., “Drowsiness Detection According to the Number of Blinking Eyes Specified from Eye Aspect Ratio Value Modification”, Proceeding of the 1st International Conference on Life, Inovation, Change and Knowledge ,2018.

[29] Taylor, M., Playsound python library https://pypi.org/project/playsound/,2015

[30] Karen, S. and Zisserman, A., “Deep Convolutional Networks for Large-Scale Image Recognition”, arXiv preprint arXiv:1409.1556 ,2014.

[31] Krizhevsky, A., Sutskever, I. and Hinton, G.E., “ImageNet Classification with Deep Convolution Neural Networks”, Advances in Neural Information Processing System 25(2),2011.

[32] Bradski, G., “The OpenCV Library. Dr. Dobb&#39”, Journal of Software Tools 25(11), pp.120-126,2000.

Author Biography

Chinh Tran received his B.S and M.S in Electrical Engineering from The Catholic University of America in 2018 and 2020, respectively, and is currently a Ph.D. student. His research interests are image motion detection, computer vision, and machine learning.

Nader Namazi (IEEE life senior member and member of Eta Kappa Nu). He is currently professor and chairman in the Department of Electrical Engineering and Computer Science. He was the Editor-in-Chief of the international Journal of Modeling and Simulation from 1994-1996. His research interests are in image motion detection and estimation, image sequence filtering/restoration, digital communications and classification. He is also recipient of the Alpha Delta Gamma Teaching Award.