Robust Drowsiness Detection for Vehicle Driver using Deep Convolutional Neural Network

A F M Saifuddin Saif[1]

Faculty of Science and Technology
American International University – Bangladesh
Dhaka, Bangladesh

Zainal Rasyid Mahayuddin[2]

Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM, Bangi, Selangor, Malaysia

Abstract—**Drowsiness detection during driving is still an unsolved research problem which needs to be addressed to reduce road accidents. Researchers have been trying to solve this problem using various methods where most of these solution lacks behind in accuracy, real-time performance, costly, complex to build, and has a higher computational cost with low frame rate. This research proposes robust method for drowsiness detection of vehicle drivers based on head pose estimation and pupil detection by extracting facial region initially. Proposed method used frame aggregation strategy in case of face region cannot be extracted in any frame due to shortcomings, i.e. light reflection, shadow. In order to improve identification under highly varying lighting conditions, proposed research used cascade of regressors cutting edge method where each regression refers estimation of facial landmarks. Proposed method used deep convolutional neural network (DCNN) for accurate pupil detection to learn non linear data pattern. In this context, challenges of varying illumination, blurring and reflections for robust pupil detection are overcome by using batch normalization for stabilizing distributions of internal activations during training phase which makes overall methodology less influenced by parameter initialization. Proposed research performed extensive experimentation where accuracy rate of 98.97% was achieved using frame rate of 35 fps which is higher comparing with previous research results. Experimental results reveal the effectiveness of the proposed methodology.**

Keywords—*Drowsiness detection; convolutional neural network; face region extraction; pupil detection*

I. INTRODUCTION

In this age of era, road accidents happens to be caused by lot of reasons where one of that reason happens to be driver feeling drowsy during driving known as fatigue driving. Previous research proposed two types of drowsiness detection system, i.e. subjective detection methods and objective detection methods. In subjective detection methods, drivers tend to determine their state of drowsiness by individual record data-sheet and own physiological response. Subjective methods are easy to build and simple in usage but lacks behind in accuracy and real-time performance. Objective detection methods use modern technologies to evaluate driver drowsiness state by extraction of driving characteristics, facial features analysis and come to a result base conclusion. However, simulation of objective detection methods are tends to be costly, complex to build, and has a higher computational cost. Objective detection methods can be divided into two categories based on detection parameters, i.e. driver

physiological parameters based detection method and driver facial features based detection method. In driver physiological parameters based detection method, normally driver response becomes slow due to drowsiness where EEG can be used to detect drowsiness state [1, 2, 3]. However, this approach tends to be costly and real time detection during driving is not robust. Driver facial features based detection methods are based on facial features using various methods, i.e. CNN based Deep Learning Model [4],Generative Adversial Networks (GAN) [5],Principal Component Analysis [6], DriCare [7] where analysis was done on pupil, eyelids and head pose to detect drowsiness. This research proposes a robust method for vehicle driver drowsiness detection using facial features based head orientation and pupil detection where frame aggregation strategy is also used to ensure facial features processing under challenging circumstances, i.e. light reflection and shadow.

Research in [4] used neural network approach for face features based detection methods and utilized facial landmarks through Convolutional Neural Network (CNN) in order to categorize driver drowsiness. Deep learning model which is small in size with high accuracy was considered as their achievement. However, in case of wearing sunglasses in lieu with bad lighting conditions, their proposed approach could not provide expected performance. Research in [5] introduced a novel framework that remedies generalization failures under represented population groups in the training dataset. They improved Convolutional Neural Network (CNN) trained for prediction by using Generative Adversarial networks (GAN) for targeted data augmentation for population groups that face similar facial attributes and highlights where the model was failing. However, due to the lack of sufficient training data for various population groups, their research requires further investigation. Research in [6] showed that there are potential social challenges regarding the application of drowsiness detection by highlighting problems in detecting dark-skinned driver faces. They focused to use unrepresentative images to train driver drowsiness detection system using vision based approach. However, their accuracy obtained decreased substantially when their approach was evaluated using more representative test set. Research in [7] proposed a system called DriCare to detect the drivers' fatigue status, such as yawning, blinking, and using video images, duration of eye closure, without equipping their bodies with devices. However, they did not show any comparison of performance with the existing research methods. Proposed method for

drowsiness detection by this research extracts facial features initially followed by head orientation and pupil detection. In addition, to learn nonlinear data pattern proposed method used aggregation strategy and integral image construction in lieu with deep convolutional neural network (DCNN) to improve efficiency. Overall contributions by this research are stated below:

- Proposed method detects drowsiness for vehicle driver in the basis of visual analysis of eye and head orientation together followed by pupil detection.

- Proposed method uses deep convolutional neural network (DCNN) to learn non-linear data pattern.

- Proposed method handles highly varying lighting conditions by using cascade of regressors cutting edge method.

- Various challenges, i.e. unfixed illumination, blurring and reflections are overcome by the proposed method using batch normalization to stabilize distributions of internal activations during training phase.

Rest of this paper is organized as follows. Section 2 demonstrates comprehensive and critical reviews in the existing research, Section 3 illustrates proposed methodology for drowsiness detection, Section 4 depicts extensive experimental validation for the proposed method and finally Section 5 presents concluding remarks.

II. BACKGROUND STUDY

Researchers used various methods for driver drowsiness detection mentioned in Fig. 1. Previously, researchers used artificial neural network (ANN) for drowsiness detection where information such as features are stored in the entire network and can learn from observing datasets and disappearances of few piece of information does not prevent network from functioning [8, 9, 14]. However, ANN requires higher computation due to parallel processing is needed which means that overall methodology will not always be robust and little suitable to use in a vehicle constantly [10, 11]. In this context, deep learning or a form of machine learning approach can be used to speed up analysis of datasets [12, 13, 15]. Research in [4] focused on the detection of drowsiness using neural network based methodology. Facial landmarks from frames captured by mobile device was detected using CNN-based machine learning approach by them and CNN-based trained deep learning model was used to detect drowsy driving behavior. However, in their research still better performance can be achieved using efficient facial features detection under bad lighting conditions. Research in [5] identified individuals with facial features by introducing a sampling strategy where the network was failing. They showed how a Generative Adversarial networks (GAN) can be used to produce training data or individuals where model is failing by generating realistic images. Their proposed approach did not rely on any meta-data or assumptions about the race or ethnicity of individuals in the datasets, which is a commonly used approach to determine algorithmic fairness or bias. However, their proposed methodology still requires that some training data be available for various population groups. A novel

visualization technique which can be assistance to identify groups of people was proposed by research in [6] where potential discrimination could arise due to the usage of Principal Component Analysis (PCA). They used PCA to produce a grid of faces sorted by similarity and combining these with a model accuracy overlay. They aimed to highlight the challenges of using unrepresentative images to train vision-based driver drowsiness detection systems. However, facial regions for the lighter skinned individuals were focused by them and failed to do so for darker skinned individuals pointed potential failure case. In addition, indication of some overfitting on more representative datasets was observed causes a decrease on more representative datasets. Research in [7], proposed a new face-tracking algorithm named Multiple Convolutional Neural Networks (CNN)-KCF (MC-KCF), which optimizes KCF algorithm. Their proposed DriCare provides three different criteria to evaluate the degree of the driver's drowsiness: blinking frequency, duration of the eyes closing, and yawning. They combined CNN with KCF algorithm to improve the performance under complex environment, such as low light. However, they did not compare effectiveness of their proposed DriCare with other existing methods.

Driving parameters in comparison with facial features and drivers physiological parameters is another widely used approach in detection of fatigue during driving. Karolinska sleepiness scale (KSS) is an ideal method for vehicle driving parameters which refers as questionnaire that depends on drivers to self-involvement and answers of drivers from the pre-set questionnaire [16, 17, 18]. Then, these answers tend to pass through KSS and results are generated from the KSS scale. Due to the simplicity of this approach to detect fatigue of vehicle driver is stated as simple approach which does not require complicated computation and processing [19, 20, 21]. However, core problem with the vehicle driving parameter in the detection of fatigue is that vehicle driving parameters often results in poor accuracy and sensitivity.
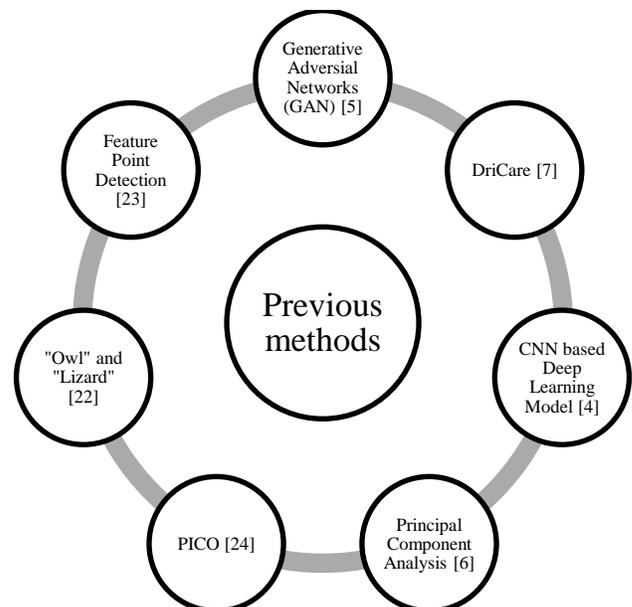

Fig. 1. Previous Research Methods for Driver Drowsiness Detection.

Research in [22] used random forest classifier for classification where they required a face and pupil under classification schemes and observed variation in accuracy among subjects before and after adding eye pose in the classification set. However, variation of performance in their proposed methodology was noticed results in unreliable validation. In this context, research in [23] experimented effects of facial sub regions to estimate accuracy where they combined eye and mouth areas. They used facial textures and landmarks extracted from image sequences and fused these features to enhance performance. However, due to lack of week feature extraction mechanism used by their research caused higher error rate. Research in [24] used facial features as parameters for drowsiness detection. They used rigid face model which enables fast pose estimation, especially for on-board computers. Their 3D face model produced more robust recognition memory. However, rigid face model cannot update shape. Besides, research in [24] absolute deviations are mainly caused by rigid model which has geometry error for different subjects resulting in higher error rate.

This research proposes robust method for drowsiness detection using facial region extraction and head orientation estimation followed by pupil detection. To ensure the robustness of the proposed methodology, aggregation strategy is used during facial region extraction to overcome various shortcomings, i.e. light reflection, shadow. In addition, to overcome the challenges of varying illumination, blurring and reflections, this research used deep convolutional neural network (DCNN) to learn non-linear data pattern and high probability values are provided to the pixels enclosed by the generated circle instead of providing pupil's coordinates directly.

III. PROPOSED METHODOLOGY

This research proposes an efficient method to reduce collisions associated with driver drowsiness. Difference between existing methods and proposed method is that proposed method detects drowsiness in the basis of visual analysis of eye state and head pose whereas in previous research methods, driver drowsiness or distraction level was determined by using eye closure or head-nodding angles. Various steps such as face detection, face alignment, pupil detection, classification and decision-making, Kalman filter are included in the proposed method mentioned in Fig. 2. All these steps for the proposed method are extensively explained.

A. Input Image

Proposed research used monocular video camera using 35 fps frame rate for input image collection. Proposed research used median filter [25, 26] to remove noise from the collected video frames. In this context, morphological processing such as resizing of the frame into 200x200 dimension, erosion and dilation were applied to ensure noise free frames to the next subsequent frames. In addition, this research also used two frame differential approaches to find the difference between frames to find the initial change between video frames.
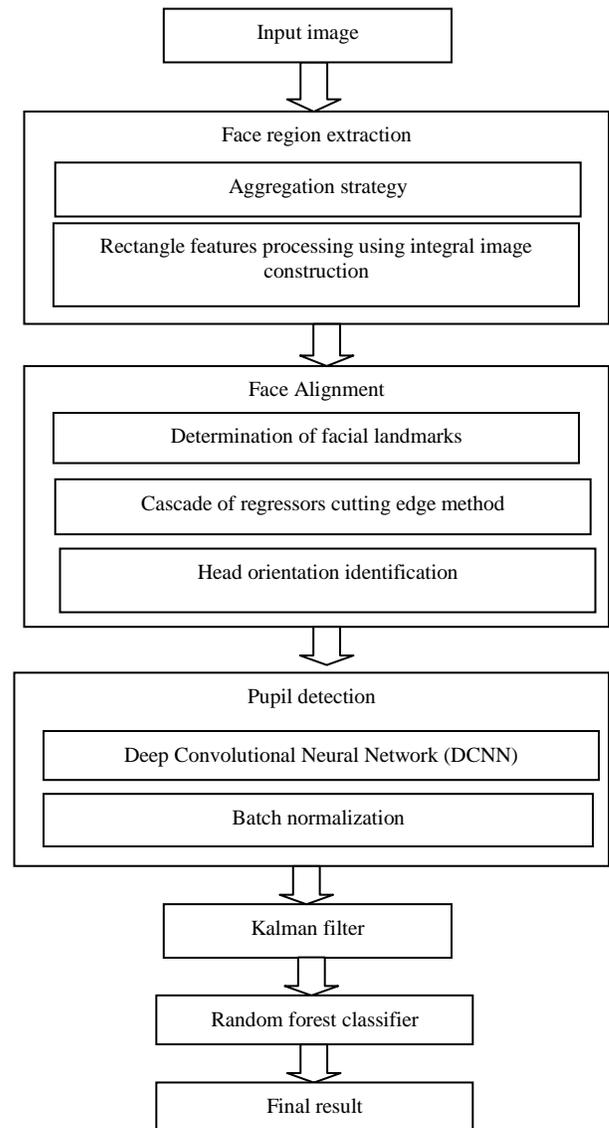

Fig. 2. Proposed Method for Drowsiness Detection.

B. Face Region Extraction

In this step, proposed research composed face detection into three more steps, i.e. implication of Viola-Jones algorithm for facial region extraction, aggregation strategy and integral image construction to process rectangle features. This research used Viola-Jones algorithm to extract facial region extraction due to maintain reliable frame rate for overall processing in every frame achieved in the previous step. In several frames, face region was not extracted due to some issues, i.e. light reflection, shadow. To overcome these shortcomings, frame aggregation strategy is used by this research. In frame aggregation strategy, frames without a detected face region are replaced by the previous frame. During face region extraction process, rectangle features are computed through intermediate representation in one pass over the original frames called as integral image. Thus, face region was extracted from the input frames achieved in the previous frames.

C. Face Alignment

After face region extraction, various facial landmarks are determined from face region components, i.e. eyes, upper edge of eyebrows, upper edge of the eyebrows, inner and outer lips, jawline, and parts in and around the eye from the location and size of individual face region. In order to improve identification under highly varying lighting conditions for monocular video frames, proposed research used cascade of regressors cutting edge method. Cascade of regression was implemented as a sequence of regressors progressively refining the estimation of poses from individual frames where each regression refers an estimation of facial landmarks. In order to localize driver gaze localization as part of facial landmarks identification, two characteristics play significant role, i.e. robust to partial occlusion and self-occlusion, Running-time of cascade of regressors cutting edge method is significantly faster than 30 fps video frame rate. These two characteristics are mapped directly to gaze location using cascade of regression to compute head orientation. Extracted facial landmarks are also be mapped to a 3D model of the head. Orientation of the head is computed by the resulting 3D–2D point correspondence.

D. Pupil Detection

Robust and accurate detection of pupil is a key for head mounted drowsiness detection after face alignment by the proposed research. Varying illumination, blurring and reflections are the main challenges for accurate pupil detection. Proposed method used deep convolutional neural network (DCNN) to obtain coordinates of the pupil center. DCNN consists of several layers which contain neurons that perform local convolutions and main variables of the model are the weights of local convolutions. Activation function such as rectified linear unit is used for DCNN network to introduce non linearity into the network in order to learn non-linear function. Proposed research used facial landmarks as the center of segmentation to be used through DCNN network. DCNN generates a mask in which pixels values provides the probability of being part of the circle centered in the pupil due to the usage of circular mask of size 25 pixels with value one was drawn in the position of pupil's center on a background of value zero. For this reason, high probability values are assigned by DCNN network to the pixels enclosed by the generated circle instead of providing pupil's coordinates directly. To overcome the challenges of varying illumination, blurring and reflections for robust pupil detection, this research used batch normalization for stabilizing distributions of internal activations during training of the model and makes model less impacted by parameter initialization. Batch normalization uses mean and variance of the mini-batch to normalize mean and variance before activation function and to estimate a moving average and variance that are used during inference.

E. Kalman Filter

Kalman filter is used to provide the best estimation of states by using measurements from various sensors in the presence of noise. Proposed method used Kalman filters to optimally estimate variables for higher accuracy rate. After pupil detection, if there is any frame loss due to noise then Kalman filter is used by the proposed method to process that frame by illuminating noise. Although, in the context of automotive drivers, face region does not change abruptly in consecutive frames, gaze point can be often still noise due to tiniest differences in the face location or endpoints of the fitted pupil ellipse cause deviation in the calculated gaze point. For this reason, estimation of the eye features and gaze point, fluctuate even when steadily fixating a single point which can even make the use of gaze information impractical. For this reason, gaze point is smoothed by Kalman filter to improve performance.

F. Classification

Finally, decision for drowsiness from the extracted and filtered landmarks is performed using random forest classifier from single feature vector by generating a set of probabilities for each class. In this context, probabilities are estimated using mean predicted class probabilities of the trees in the forest where class probability of a single tree is the fraction of samples of the same class in the tree. Class with highest probability is the one that is assigned to the image as the "decision". In this context, ratio of the highest probability to the second highest probability is referred to as "confidence" of the decision. Proposed research fixed threshold of 10 based on various trials and depending on the facial landmark features with head pose movement and designated classes based on drowsiness behavior. Any decision with confidence more than threshold are accepted and others are ignored.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Hardware and Software Set up

Hardware and software set up are initial core part for validating proposed method before starting experimentation [27, 28]. Collection of datasets was done in real time during hardware set up followed by software set up. Usage the type of camera and CPU were fixed in the hardware set up stage where proposed method used monocular camera. Monocular camera has some advantages comparing with binocular and webcams. Monocular camera provides single eyepiece with good resolution during capturing frames from real time scene [29, 30]. In addition, this type of camera consists of night vision so low light issue or capturing frames during night time does not impact in the datasets used by this research. Besides, low cost and robustness of the monocular camera in lieu with reducing computational cost by using one camera instead of using multiple web cameras encourages proposed research to use monocular camera. Proposed research used Intel Core i7 8th generation to serve the purpose of processing training datasets faster. Python programming language was used during experimentation to validate the proposed method.

B. Datasets

Proposed method was validated using iBUG 300 W dataset which consists of 300 indoor and 300 outdoor images [22]. iBUG 300 W dataset contains diverse pool of images in various perspectives, i.e. naturalistic, unconstrained face images, different variations such as pose, expression, background, occlusion, resolution, various expression such neutral, surprise, squint, smile, disgust, scream. In addition, various images related to various expressions were collected

from various sources, i.e. party, conference, protest, sport, celebrities. Besides, re-annotation was performed in iBUG 300 W datasets using semi-supervised approach to improve accuracy.

C. Experimental Results

This research used various performance metrics, i.e. accuracy, error rate, frame rate as frame per second (fps) to validate proposed method [31, 32, 33]. Accuracy of the proposed method is calculated using (1) mentioned below:

$$K_a = \frac{k_{nt}}{k_{nt}+k_{rp}} \qquad (1)$$

Here, k_{nt} denotes distance of nose tip and k_{rp} denotes distance of right pupil from their average position of the background model. Both distances are in the range of $[0, \sqrt{2}]$ due to normalization of the features. Value of K_a, 0 indicates a shift in gaze involves only in eyes and value of K_a, 1 indicates a shift in gaze involves only head. Measurement of drowsiness for each subject is computed by averaging the result of K_a. This research categorized subjects into three groups, i.e. gaze in eye, gaze in head and mixed gaze based on the value of K_a using adaptive threshold mentioned in (2) [34, 35].

$$\theta = \psi - \frac{\psi \times (\log_2(K_a)+1)}{100} \qquad (2)$$

Here, K_a denotes total number of frames, mean value of pictorial intensities is denoted as ψ in a video frame. Threshold value is denoted as θ.

Proposed method achieved accuracy rate of 98.97% which is higher than previous research methods. Proposed method received minimum error rate of 1.03% using frame rate of 35 fps shown in Table I.

D. Comparision with Previous Research Results

Experimental results of the proposed method are compared with previous research results based on accuracy, error rate and frame rate shown in Table II, Table III and Table IV respectively. Research in [4] received accuracy rate of 83% using CNN-based trained Deep Learning model and used frame rate of 30 fps. In the presences of low light condition, their research could not provide satisfactory validation. Research in [5] received accuracy rate of 98.01% using Generative Adversarial Networks (GAN). However, lack of training data initiates the need of further investigation to validate their overall methodology. Research in [6] received accuracy rate of 98.7% using Principal Component Analysis (PCA) to deal with unrepresentative models. However, their research could not show expected accuracy rate when tested on more representative test dataset. Research in [7] received accuracy rate of 92% using their proposed DriCare. They used frame rate of 18 fps when environment was bright and 16 fps when environment was dark. However, they did not compare DriCare with existing research methods. Research in [24] received accuracy rate of 88.41% and 77.8% for day datasets and night datasets separately using PICO (Optimized).

However, PICO (Optimized) received error rate of 11.59%.Viola-Jones (OpenCV) based face detection methodology in the same research received accuracy rate of 78.43% for daytime dataset and 76.01% for night time datasets using 9 fps frame rate. However, higher error rate of 11.9% and 13.99% were received for day time and night time datasets respectively which are higher comparing with the proposed method by this research. Research in [36] received accuracy rate of 96.7% using 24fps frame rate by adopting face landmark. However, accuracy rate in the research fluctuates when temporal resolution was higher than 6 fps. Their research could not deal with interferences constraint which causes low performances comparing with the proposed method by this research. In addition, they used texture information which decreased average accuracy due to variations in illumination and skin colour. In this context, proposed method by this research used information, i.e. head and eye using 35 fps frame rate which causes higher accuracy than research in [36].Research in [23] achieved accuracy rate of 95% using feature point extraction. They received error rate less than 5% by reducing computational complexity effectively. However, their method provided efficient result using 25 fps frame rate only whereas proposed method by this research has the ability to perform with 35 fps frame rate. In addition, proposed method received lower error rate of 1.03% comparing with research in [23]. Research in [22] received accuracy rate of 94.6% using 30 fps frame rate which also indicates low performance comparing with the proposed method by this research. In addition, variation in accuracy among subjects before and after adding eye pose to the classification set, research in [22] requires further investigation.

TABLE I. EXPERIMENTAL RESULTS USING PROPOSED METHOD

Accuracy (%)	Error Rate (%)	Frame Rate (FPS)
98.97	1.03	35

TABLE II. COMPARISON WITH PREVIOUS RESEARCH METHODS BASED ON ACCURACY

Methods		Accuracy
Proposed method		98.97%
CNN based Deep Learning model [4]		83%
Generative Adversarial networks (GAN) [5]		98.01%
Principal Component Analysis (PCA) [6]		98.7%
DriCare [7]		92%
Feature point detection [23]		95%
Facial landmark location [36]		96.7%
"owl" and "lizard" [22]		94.6%
PICO (Optimized) [24]	Day time	88.41%
	Night time	77.85%
Viola-Jones (OpenCV)[24]	Day time	78.43%
	Night time	76.01%

TABLE III.   COMPARISON WITH PREVIOUS RESEARCH METHODS BASED ON ERROR RATE

Methods		Error Rate
Proposed method		1.03 %
PICO (Optimized) [24]		11.59%
Viola-Jones (OpenCV)[24]	Day time	11.9%
	Night time	13.99%
Feature point detection [23]		Less than 5%

TABLE IV.   COMPARISON WITH PREVIOUS RESEARCH METHODS BASED ON FRAME RATE

Methods	Frame Rate
Proposed method	35 fps
CNN based Deep Learning model [4]	30 fps
DriCare [7]	18 fps
Feature point detection [23]	25 fps
Facial landmark location [36]	24 fps
"owl" and "lizard" [21]	30 fps
Viola-Jones (OpenCV)[24]	9 fps

E. Analysis and Discussion

Although previous research attempted to achieve effective methodology for drowsiness detection, various concerns are still existed in terms with validating their methodology. Research in [4] used CNN based machine learning approach and achieved accuracy rate of 83% using 30 fps. They were able to detect facial landmarks to pass through CNN based deep learning model. However, in the presence of low light condition and obstructions due to wear glass by the subjects, their proposed methodology, could not provide expected validation results. Proposed method by this research used rich datasets where glasses or low light conditional scenario were overcome indicating by higher accuracy rate of 98.97% with higher frame rate of 35 fps than research in [4] shown in Fig. 3. Research in [5] used Generative Adversarial Networks (GAN) and introduced a sampling strategy to identify individuals with facial features where the network was failing. Although, they achieved accuracy rate of 98.01%, due to lack of population groups during validation, their methodology demands for further investigation to improve performance. In this context, research in [6] achieved accuracy rate of 98.7% by using Principal Component Analysis (PCA) where similarity based sorting was done by producing a grid of faces and combining these with a model accuracy overlay. However, Principal Component Analysis (PCA) method is mostly known as dimension reduction method which mainly depends of efficient extraction of features [13, 25]. In addition, in their research, accuracy decreased substantially when their proposed approach was validated using more representative test set. Accuracy of the proposed method by this research was not decreased during any stage of experimentation due to the usage of geometric structure identification of human face in lieu with that proposed method received lower error rate comparing with previous research results. Research in [7] proposed DriCare to estimate different

criteria about the degree of the driver's drowsiness, i.e. blinking frequency, duration of eyes closure. They achieved accuracy rate of 92% using two different frame rates, i.e.18fps when the environment was bright, 16 fps when the environment was dark. In both context of accuracy rate and frame rate, proposed method achieved higher accuracy rate of 98.97% using higher frame rate of 35 fps. Research in [23] assessed effects of facial sub regions to estimate accuracy where they combined eye and mouth areas and caused accuracy rate of 95% in lieu with error rate of less than 5% using 25 fps shown in Fig. 4 and Fig. 5. They used facial texture and landmarks extracted from image sequences and fused these features to enhance the performance causes higher error rate than proposed method by this research. In this context, proposed method used pupil and head position to detect drowsiness causes better accuracy rate in lieu with frame rate of 35 fps using 1.03% error rate which are better than research in [23].
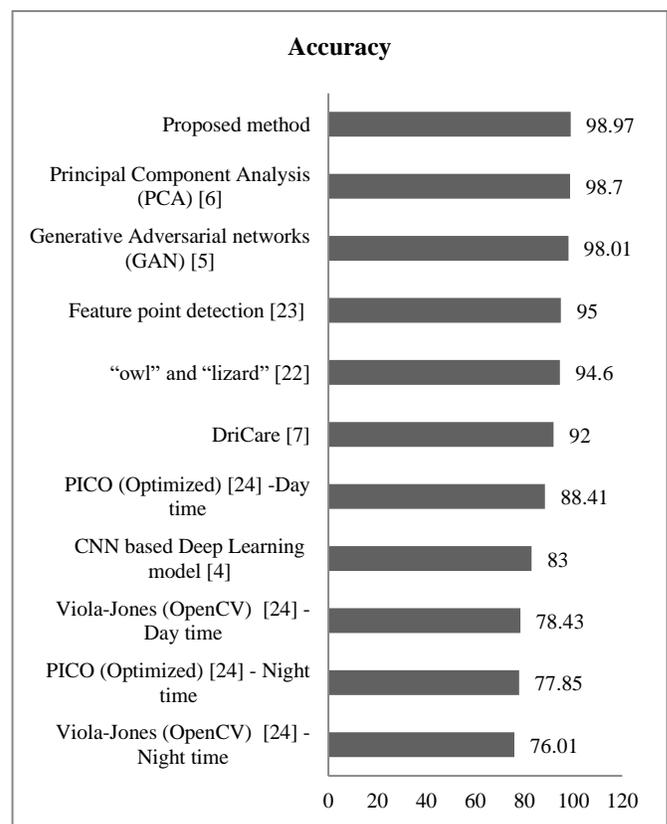

Fig. 3.   Comparison among Proposed Method and Previous Research based on Accuracy.

Research in [22] required a face and pupil under classification schemes where they observed variation in accuracy among subjects before and after adding eye pose to the classification set. Although they received accuracy rate of 94.6% using 30 fps frame rate, their accuracy rate is not reliable enough due to the variation in accuracy which drops to below 80% and even to as low as 40%.In this context, proposed research used pupil and head position detection to detect drowsiness using 35 fps with higher accuracy rate of 98.97% than research in [22]. Research in [24], absolute deviations are mainly caused by the rigid face model, which

has geometry error for different subjects, and cannot update its shape like the flexible model or adaptive model. For this reason, research in [24] received higher error rate comparing with the proposed method by this research. In this context, proposed method detects facial regions of a driver when an image sequence is captured in lieu with applying Viola-Jones algorithm for facial region extraction. Proposed method used Viola-Jones algorithm for facial region extraction with higher frame rate processing perspectives which causes lower error rate at the same time.
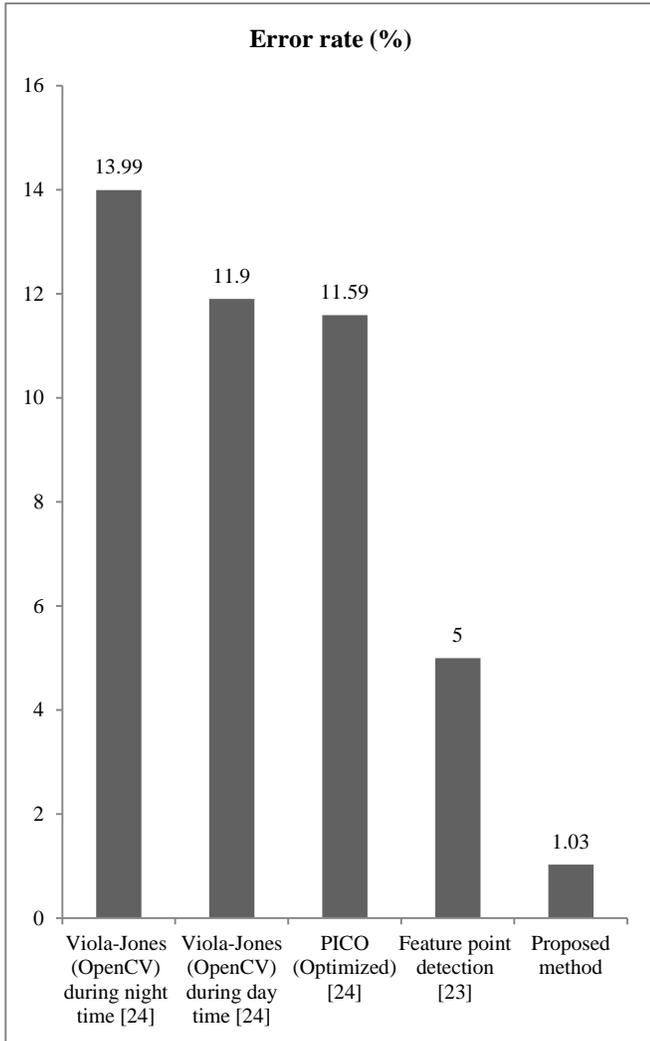
Error rate (%)

Fig. 4. Comparison among Proposed Method and Previous Research based on Error Rate.
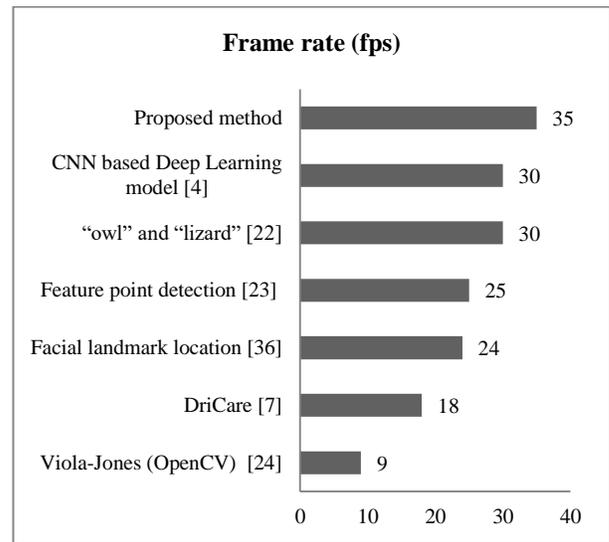
Frame rate (fps)

Fig. 5. Comparison among Proposed Method and Previous Research based on Frame Rate.

V. Conclusion

Proposed method detects driver drowsiness by extracting facial features initially and then performs face alignment following by pupil detection and classification by random forest classifier. During facial features extractions proposed method used aggregation strategy and integral image construction to process rectangle features in case face region cannot be extracted due to some issues, i.e. light reflection, shadow in input frames. In face alignment phase, proposed method used cascade of regressors cutting edge method in order to improve identification of facial landmarks under highly varying lighting conditions for video frames. Later, in pupil detection step, proposed research used deep convolutional neural network (DCNN) for accurate pupil detection for nonlinear data pattern where proposed method used facial landmarks as the center of segmentation to be used through DCNN network. Although, in the context of automotive drivers, face region does not change abruptly in consecutive frames, gaze point is smoothed by Kalman filter to remove noise due to tiniest differences in the face location or endpoints of the fitted pupil ellipse cause deviation in the calculated gaze point and thus improve performance. In this context, appropriate features selection during facial features extraction in lieu with computational time measurement to compare with previous research is intended to investigate in future. Experimental results for the proposed method reveal higher efficiency comparing with previous research results in terms with accuracy rate, error rate and frame rate. Proposed method achieved accuracy rate of 98.97% with error rate of 1.03% using frame rate of 35 fps. Performance of the proposed method reveals the potentiality to impact significantly to reduce road accidents.

REFERENCES

[1] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of driver vigilance level using EEG signals and driving contexts," IEEE Transactions on Reliability, vol. 67, pp. 370-380, 2017.

[2] Z. Zhang, D. Luo, Y. Rasim, Y. Li, G. Meng, J. Xu, and C. Wang, "A vehicle active safety model: Vehicle speed control based on driver vigilance detection using wearable EEG and sparse representation," Sensors, vol. 16, p. 242, 2016.

[3] Y. Liang, W. J. Horrey, M. E. Howard, M. L. Lee, C. Anderson, M. S. Shreeve, C. S. O'Brien, and C. A. Czeisler, "Prediction of drowsiness events in night shift workers during morning driving," Accident Analysis & Prevention, vol. 126, pp. 105-114, 2019.

[4] R. Jabbar, M. Shinoy, M. Kharbeche, K. Al-Khalifa, M. Krichen, and K. Barkaoui, "Driver drowsiness detection model using convolutional neural networks techniques for android application," in 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), 2020, pp. 237-242.

[5] M. Ngxande, J.-R. Tapamo, and M. Burke, "Bias Remediation in Driver Drowsiness Detection Systems Using Generative Adversarial Networks," IEEE Access, vol. 8, pp. 55592-55601, 2020.

[6] M. Ngxande, J.-R. Tapamo, and M. Burke, "Detecting inter-sectional accuracy differences in driver drowsiness detection algorithms," in 2020 International SAUPEC/RobMech/PRASA Conference, 2020, pp. 1-6.

[7] W. Deng and R. Wu, "Real-time driver-drowsiness detection system using facial features," IEEE Access, vol. 7, pp. 118727-118738, 2019.

[8] C. J. de Naurois, C. Bourdin, A. Stratulat, E. Diaz, and J.-L. Vercher, "Detection and prediction of driver drowsiness using artificial neural network models," Accident Analysis & Prevention, vol. 126, pp. 95-104, 2019.

[9] L. Chukoskie, S. Guo, E. Ho, Y. Zheng, Q. Chen, V. Meng, J. Cao, N. Devgan, S. Wu, and P. C. Cosman, "Quantifying gaze behavior during real-world interactions using automated object, face, and fixation detection," IEEE Transactions on Cognitive and Developmental Systems, vol. 10, pp. 1143-1152, 2018.

[10] Z. R. Mahayuddin and A. F. M. S. Saif, "A Comprehensive Review Towards Segmentation and Detection of Cancer Cell and Tumor for Dynamic 3D Reconstruction," Asia-Pacific Journal of Information Technology and Multimedia, vol. 9, pp. 28-39, 2020.

[11] Z. R. Mahayuddin and A. S. Saif, "A Comprehensive Review Towards Appropriate Feature Selection for Moving Object Detection Using Aerial Images," in International Visual Informatics Conference, 2019, pp. 227-236.

[12] Z. R. Mahayuddin and A. S. Saif, "A Comparative Study of Three Corner Feature based Moving Object Detection Using Aerial Images," Malaysian Journal of Computer Science, pp. 25-33, 2019.

[13] A. S. Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Moment feature based fast feature extraction algorithm for moving object detection using aerial images," PloS one, vol. 10, p. e0126212, 2015.

[14] M. H. Alkinani, W. Z. Khan, and Q. Arshad, "Detecting Human Driver Inattentive and Aggressive Driving Behavior Using Deep Learning: Recent Advances, Requirements and Open Challenges," IEEE Access, vol. 8, pp. 105008-105030, 2020.

[15] M. Dua, R. Singla, S. Raj, and A. Jangra, "Deep CNN models-based ensemble approach to driver drowsiness detection," Neural Computing and Applications, pp. 1-14, 2020.

[16] T. Wilaiprasitporn and T. Yagi, "Feasibility study of drowsiness detection using hybrid brain-computer interface," in Proceedings of the international Convention on Rehabilitation Engineering & Assistive Technology, 2016, pp. 1-4.

[17] B.-L. Lee, B.-G. Lee, and W.-Y. Chung, "Standalone wearable driver drowsiness detection system in a smartwatch," IEEE Sensors Journal, vol. 16, pp. 5444-5451, 2016.

[18] C. B. S. Maior, M. J. das Chagas Moura, J. M. M. Santana, and I. D. Lins, "Real-time Classification for Autonomous Drowsiness Detection Using Eye Aspect Ratio," Expert Systems with Applications, p. 113505, 2020.

[19] X. Zhang, X. Wang, X. Yang, C. Xu, X. Zhu, and J. Wei, "Driver drowsiness detection using mixed-effect ordered logit model considering time cumulative effect," Analytic methods in accident research, vol. 26, p. 100114, 2020.

[20] T. Kundinger, A. Riener, N. Sofra, and K. Weigl, "Driver drowsiness in automated and manual driving: insights from a test track study," in Proceedings of the 25th International Conference on Intelligent User Interfaces, 2020, pp. 369-379.

[21] K. Mahajan and N. R. Velaga, "Effects of partial sleep deprivation on braking response of drivers in hazard scenarios," Accident Analysis & Prevention, vol. 142, p. 105545, 2020.

[22] L. Fridman, J. Lee, B. Reimer, and T. Victor, "'Owl'and 'Lizard': Patterns of head pose and eye pose in driver gaze classification," IET Computer Vision, vol. 10, pp. 308-314, 2016.

[23] C. Meng and X. Zhao, "Webcam-based eye movement analysis using CNN," IEEE Access, vol. 5, pp. 19581-19587, 2017.

[24] C. Yin and X. Yang, "Real-time head pose estimation for driver assistance system using low-cost on-board computer," in Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1, 2016, pp. 43-46.

[25] Z. R. Mahayuddin and A. F. M. S. Saif, "Efficient Hand Gesture Recognition Using Modified Extrusion Method based on Augmented Reality," TEST Engineering and Management, vol. 83, pp. 4020-4027, 2020.

[26] Z. R. Mahayuddin and A. F. M. S. Saif, "Augmented Reality based Ar Alphabets Towards Improved Learning Process in Primary Education System," Journal of Critical Reviews, vol. 7, 2020.

[27] Z. R. Mahayuddin, A. S. Saif, and A. S. Prabuwono, "Efficiency measurement of various denoise techniques for moving object detection using aerial images," in 2015 International Conference on Electrical Engineering and Informatics (ICEEI), 2015, pp. 161-165.

[28] A. Saif and Z. R. Mahayuddin, "Moving Object Segmentation Using Various Features from Aerial Images: A Review," Advanced Science Letters, vol. 24, pp. 961-965, 2018.

[29] A. Saif, A. Prabuwono, and Z. Mahayuddin, "Adaptive long term motion pattern analysis for moving object detection using UAV aerial images," International Journal of Information System and Engineering, vol. 1, pp. 50-59, 2013.

[30] A. Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Moving object detection using dynamic motion modelling from UAV aerial images," The Scientific World Journal, vol. 2014, 2014.

[31] A. S. Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Adaptive motion pattern analysis for machine vision based moving detection from UAV aerial images," in International Visual Informatics Conference, 2013, pp. 104-114.

[32] A. S. Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Real time vision based object detection from UAV aerial images: a conceptual framework," in FIRA RoboWorld Congress, 2013, pp. 265-274.

[33] A. S. Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Motion analysis for moving object detection from UAV aerial images: A review," in 2014 International Conference on Informatics, Electronics & Vision (ICIEV), 2014, pp. 1-6.

[34] A. S. Saif, A. S. Prabuwono, Z. R. Mahayuddin, and T. Mantoro, "Vision-based human face recognition using extended principal component analysis," International Journal of Mobile Computing and Multimedia Communications (IJMCMC), vol. 5, pp. 82-94, 2013.

[35] A. S. Saif, A. S. Prabuwono, Z. R. Mahayuddin, and H. T. Himawan, "A review of machine vision based on moving objects: object detection from UAV aerial images," International Journal of Advancements in Computing Technology, vol. 5, p. 57, 2013.

[36] L. Zhao, Z. Wang, X. Wang, and Q. Liu, "Driver drowsiness detection using facial dynamic fusion information and a DBN," IET Intelligent Transport Systems, vol. 12, pp. 127-133, 2017.