SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ DERGİSİ

Sakarya University Journal of Science
SAUJS

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University | http://www.saujs.sakarya.edu.tr/

Title: Temporal Analysis based Driver Drowsiness Detection System using Deep Learning Approaches

Authors: Furkan KUMRAL, Ayhan KÜÇÜKMANİSA

Recieved: 2022-02-15 00:00:00

Accepted: 2022-05-31 00:00:00

Article Type: Research Article

Volume: 26
Issue: 4
Month: August
Year: 2022
Pages: 710-719

How to cite
Furkan KUMRAL, Ayhan KÜÇÜKMANİSA; (2022), Temporal Analysis based Driver Drowsiness Detection System using Deep Learning Approaches. Sakarya University Journal of Science, 26(4), 710-719, DOI: 10.16984/saufenbilder.1071863
Access link
http://www.saujs.sakarya.edu.tr/en/pub/issue/72361/1071863

Sakarya University Journal of Science 26(4), 710-719, 2022

Temporal Analysis Based Driver Drowsiness Detection System Using Deep Learning Approaches

Furkan KUMRAL[1], Ayhan KÜÇÜKMANİSA*[1]

Abstract

With the development of technology, artificial intelligence comes into our lives more and also comes up as a solution to many problems. Recently, deep learning approaches have been bringing fast and highly accurate solutions to problems. In this work, within the scope of Advanced Driver Assistance Systems (ADAS), deep learning based driver drowsiness detection system is proposed. First, face regions of drivers are detected using SSD MobileNet object detection method. The aim is to detect the eye, mouth and head positions of the drivers from this face region and to make a situation estimation with the combinations of these detected objects which are "normal", "drowsy" and "danger". The proposed approach examines the driver's behaviour over a certain period of time for making a decision, rather than a one-time eye closure or yawning decision. The detected eye, mouth and head positions are monitored and recorded over a period of time. Finally, these merged patterns are classified with Convolutional Neural Networks (CNN). Experimental results show that the performance of proposed novel CNN approach outperforms existing approaches in literature.

Keywords: Driver drowsiness, deep learning, ADAS.

1. INTRODUCTION

According to the statistics of the Traffic Directorate of the General Directorate of Security [1], there were a total of 108,171 traffic accidents until April 2021 in Turkey, and it was determined that 88.16% of these traffic accidents were caused by drivers. Drivers who make mistakes such as not following the lane, not adapting the vehicle speed to the conditions required by the road, weather and traffic, and hitting the vehicle in front with an uncontrolled speed from the rear constitute the reason for the high percentage of these accident statistics. In 2019, there were a total of 1168144 traffic accidents and 174896 of them included 5473 deaths and 283234 injuries, in Turkey [2]. Another statistic is that, there were 33,244 fatal motor vehicle crashes in the United States in 2019 in which 36,096 deaths occurred. This resulted in 11.0 deaths per 100,000 people and 1.11 deaths per 100 million miles traveled. The fatality rate per 100,000 people ranged from 3.3 in the District of Columbia to 25.4 in

* Corresponding ayhan.kucukmanisa@gmail.com
[1] Kocaeli University
E-mail: kumralf@hotmail.com
ORCID: https://orcid.org/0000-0002-1886-1250, https://orcid.org/0000-0001-6762-713X

Wyoming. The death rate per 100 million miles traveled ranged from 0.51 in Massachusetts to 1.73 in South Carolina [3].

The excessive tiredness of the drivers increases the causes of the above-mentioned accidents. Drowsiness and inattention cause the drivers to not be able to control the vehicle, weaken his reflexes during maneuvering, and not be able to perceive the distance between him and the vehicle in front of him. For this reason, the development of a system that monitors the driver's attention and drowsiness, and helps the driver to concentrate in dangerous situations, is of great importance in order to prevent the driver from causing an accident due to drowsiness during long-distance and/or night trips. Systems that warn drivers and try to prevent accidents are called Advanced Driver Support System (ADAS).

Although the driver drowsiness systems which is one of the ADAS are still subject of research, there are systems used by well-known companies today [4]. In Ford’s Driver Alert system, whether the driver stays in the lane is monitored by a camera, if the driver turns the steering wheel out of the lane, a vibration occurs in the steering wheel and the vehicle moves the steering wheel to automatically return to the lane.

Mercedes' Attention Assist fatigue monitoring system uses a steering wheel mounted sensor. This sensor can detect speed and movements of wheel and learn the behavior of the driver. This system is available in all Mercedes top segment vehicles. The company has determined that the steering wheel provides the best fatigue detection system, as tired drivers make very rapid and very small movements while steering. When the system detects an unusual steering wheel movement, it evaluates 70 other parameters such as how long the driver has been behind the wheel, the hours of travel and so on. If it is detected that the driver is drowsy, the warning “fatigue detected” appears on the instrument panel. There is also a dashboard in the car that displays the time elapsed since the driver's last stop and the driver's level of attention. When fatigue is detected, drivers receive a warning with an alarm and a search for rest areas is initiated.

Nissan's Driver Fatigue System, like BMW's Attention Assist System in the 6 and 7 series, monitors the driver's behavior during the journey via the steering wheel and alerts a drowsy drive. BMW also allows drivers to customize fatigue system related alerts; When set to “sensitive”, the driver assistance system is activated one hour after the start of the journey, and 90 minutes after starting to drive when set to “normal”.

It is difficult to create a detection system that could only identify vehicle behaviors associated with sleep-deprived drivers and fatigue. Drivers can relinquish control of the steering wheel to take care of their baby or pick up a fallen item. Also, winds hitting the car can trick the system into thinking the driver is drowsy. When a system’s sensors give too many false warnings, most drivers choose to turn them off.

Today, artificial intelligence-based approaches are used in driver drowsiness systems. In this preference, the effect of the performance increase provided by deep learning which is an artificial intelligence method is important.

The main contributions of this paper are summarized as follows:

- Temporal analysis based driver drowsiness detection approach that is different compared to most methods in the literature.
- Comparing the performance of the proposed method with image processing and deep learning based methods in comprehensive analysis.
- Suitable for low complexity embedded platforms with its low computational load.

2. RELATED WORK

When the studies in the literature are examined, it has been seen that the methods used in Advanced Driver Support Systems are considered as two separate approaches, in-vehicle and out-of-vehicle.

Systems using the out-of-vehicle monitoring approach control the factors outside the vehicle, such as whether the vehicle leaves the lane, how much it deviates, and whether the driver can keep forward vehicle distance. In-vehicle monitoring systems generally control the driver's driving status using computer vision methods. These methods are also diversified within themselves. EAR (Eye Aspect Ratio) and PERCLOS (PERcentage of eye CLOSure) are the most basic known fatigue detection methods [5]. In EAR, 6 points are determined around the eye image and these points follow the eye's circumference in the positions when the eye is closed and opened. When the eye is open and closed, the values of the mathematical operation performed with these 6 points are different from each other. In this way, it is detected whether the eye is open or closed. In these algorithms, a counter starts counting during the time the eyes are closed and when it exceeds a certain threshold value, the alarm system is activated. When the eye is opened, the number in the counter decreases and the alarm system is disabled because it falls below the threshold value.

In [6], which is basically use similar approach to the proposed method, eye and face detection are considered as an object detection approach. Eyes are labeled and trained as open and closed. Fatigue decision is determined by determining a threshold value, as mentioned in previous studies, and depending on whether the duration of eye closure exceeds this threshold value. This threshold value is determined as 40% of the frame rate of the image captured from the camera. For example, if camera is 30 FPS (Frame Per Second), the threshold value is 12. If blindfold is detected in at least 12 of the 30 frames, the alarm is activated.

Apart from these methods, there are also systems in which in-vehicle and out-of-vehicle approaches are used together. Systems have been proposed to detect signs of fatigue such as closing eyes, tilting the head, and yawning, and also checking whether there is any deviation from the lane by road observation [7]. In these systems, if there is a lane violation and there is a fatigue detection in the facial expressions of the driver simultaneously, the system gives a warning.

Recent study [8] presents an improved drowsiness detection system based on CNN-based approach. The main objective of this method is to render a system that is lightweight to be implemented in embedded systems while maintaining and achieving high performance. The system is able to detect facial landmarks from images captured on a mobile device and inference it with a deep CNN model to detect drowsy driving behavior. Another study [9] benefits from image processing methods to detect driver drowsiness. These methods are calculating the percentage of white in eyes, detection of skin surrounding the eye and Hough transform to detect the iris. Calculation of distance between the eyelids yields the best result as a final step. In [10], a system is proposed based on monitoring the driver's face of using a camera and using image processing methods to obtain some physical indicators, such as closing the eyes, blinking speed and PERCLOS, and detection of yawning, gaze direction. The driver's drowsiness is manifested as a result of features of fatigue on the driver's face. [11] propose a condition-adaptive representation learning method for efficient driver drowsiness detection which is invariant to various driving conditions containing a driving time such as day and night and a driver's appearance. The spatio-temporal representation is extracted and merged with the vectors that represent the scene understanding results using the feature fusion method based on the tensor product approach. These problems are effectively modelled using 3D-DCNN (3 Dimensional-Deep Convolutional Neural Network) and fully connected neural network.

3. PROPOSED METHOD

In this work, a deep learning based in-vehicle drowsiness detection approach is proposed. Detection of driver drowsiness is considered as an object detection problem. The eye, mouth and head positions on the image which captured from a camera towards to the driver, are determined by object detection algorithm.

The states of the detected eye, mouth and head position are recorded as a series. For example, closed eyes and yawning indicate drowsiness, open eyes and closed mouths indicate normal conditions, and tilting the head forward indicates danger. However, these decisions are not made after one-time determinations. For example, it does not mean that a driver sleeps every time he closes his eyes and yawns every time he opens his mouth. In addition, trying to understand that a person is tired simply by closing the eyes or yawning can lead to making wrong decisions. In order to prevent this, a novel approach is developed in this work. First, frame-by-frame object detections are made on the incoming frames. These objects are detected as open eyes, closed eyes, sunglasses, closed mouth, talking, yawning, and tilting the head. The driver is monitored for 50 frames (approximately 2 seconds) in the videos with approximately 25 FPS and the detected driver movements are recorded as an image. Then, an image classification approach is applied on these images with convolutional neural network. By this way, considering detected eye, mouth and head situations normal, drowsy or dangerous situation are detected. Flowchart of the proposed method is given in Figure 1.

3.1. Dataset

In this work YawDD [12], NTHU-DDD [13] and custom dataset are used in training and test stages. YawDD is a dataset of videos of drivers (men and women of different ethnicities, with and without glasses) in a real car, talking, singing, silent and yawning, recorded by a dash camera. Algorithms for yawn detection are used to develop and test models, as well as to recognize and track face and mouth. Videos shot in natural and variable lighting conditions come in two sets: In the first set, the camera is placed under the car's front mirror. This set contains 322 videos, each for a different situation: there are 3 or 4 videos of each topic: normal driving, talking or singing while driving, and yawning while driving. In the second set, the camera is installed on the driver control panel. This set includes 29 videos, one for each topic, featuring quiet driving, talking driving and driving while yawning. Certain frames are selected from these 29 videos and labeled. The number of images annotated in the YawDD dataset is 1675. Example labeling of an image from this dataset is shown in Figure 2. In this figure, it is seen that the labels 'open eye' and 'talking' are given. (The act of laughing is included in the 'talking' class).

Since YawDD is not enough with 7 different classes (open eye, closed eye, sunglasses, closed mouth, speech, yawning, head tilt) to be used in this work, another custom data set is created. In this custom data set, various videos are collected from 21 different people in their normal and

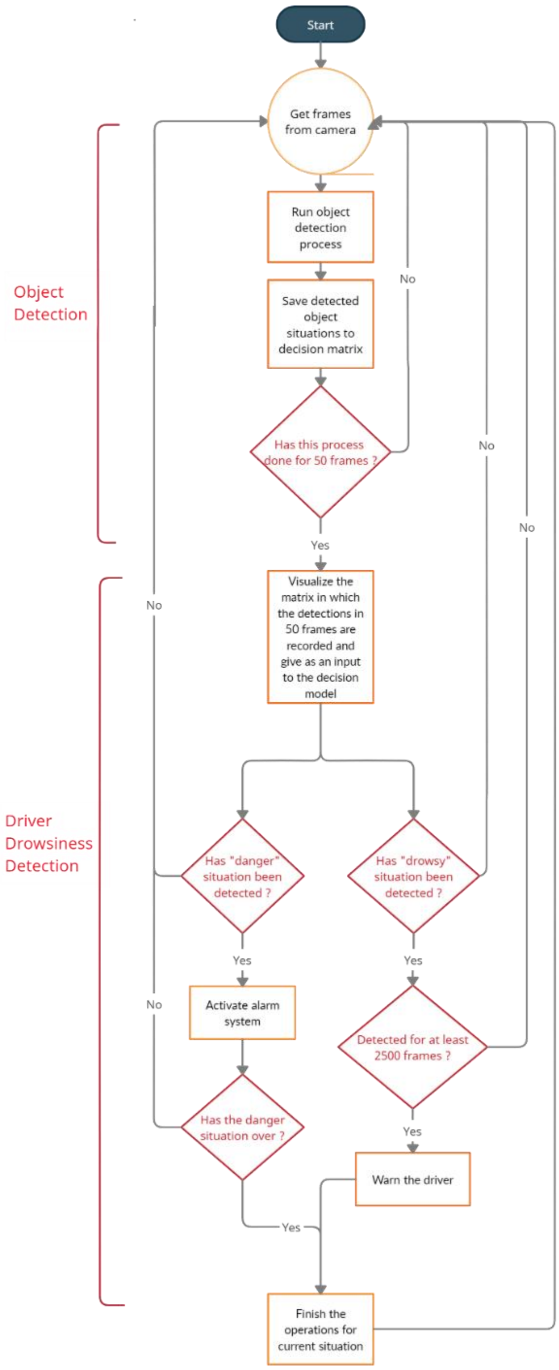


Figure 1 Flowchart of the proposed algorithm

sunglasses state. There are 45 videos in total. Certain frames are selected from these videos and labeled. An example labeling is given in Figure 2. In this figure, it is seen that the labels 'lean forward' and 'sunglasses' are given. The number of photos annotated in this custom dataset is 4504. There are 6179 annotated images for object detection model training.

3.2. Deep Learning based Object Detection

Object detection approach is used to detect the instant status of the driver in the images taken from the camera. In recent years, deep learning-based approaches have been used for the object detection function, which has achieved significant performance. Since real-time operation is an important criterion of the proposed method, the SSD MobileNet approach is preferred considering the balance of speed and performance. In addition, since this model does not have a complex neural network architecture and is a lightweight model, it is very convenient to integrate it into embedded systems. SSD MobileNet V2 FPNLite 640×640 is used as pre-trained model. There are 7 object class as “open_eye”, “close_eye”, “talking”, “closed_mouth”, “yawning”, “sunglasses” and “head_down”. Example annotations from YawDD and custom dataset are shown in Figure 2.

3.3. Driver Drowsiness Detection

After the object detection step, a novel driver drowsiness detection method is started. Flowchart of the driver drowsiness detection method is given in Figure 1 (mentioned with Driver Drowsiness Detection part).

In this this work, it is aimed not to make the decision of the driver's drowsiness status based on a single action, but to make it from the driver's general state and movements, and also to carry out this process on an ongoing basis. Thus, it is prevented to detect a drowsiness state every time the driver yawns or closes his/her eyes. The basis of this algorithm is a matrix registration system. Detected eye, mouth and head state results are recorded in a matrix. Number values between “0-6” comes as output from the object detection model.

0 : Open eye

1 : Closed eye

2 : Talking

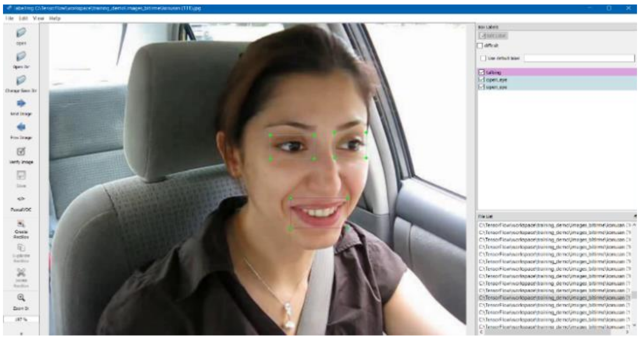
3: Closed mouth

4 : Yawn

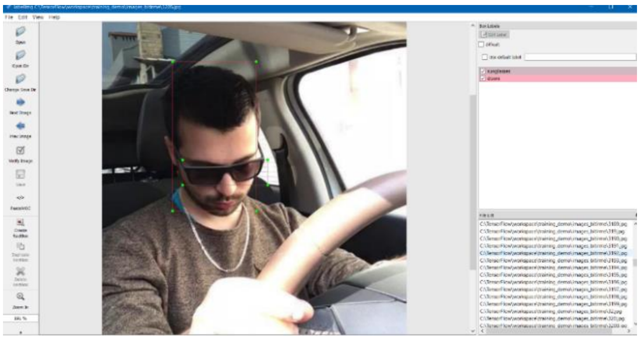
5 : Sunglasses

6 : Head tilt

This detected label used as a matrix index. If one of these states detected, index of these state value is become “1”. For example, if “Open eye” and “Closed mouth” states are detected for a driver, these detections are recorded in a list as [1,0,0,1,0,0,0]. Table 1 gives the instantaneous information about the driver drowsiness by these frame-based situations.



(a)



(b)

Figure 2 Annotations examples from datasets (a) YawDD dataset (b) Custom dataset

Table 1 Training parameters

OE	CE	T	CM	Y	S	HT	DS
1	0	1	0	0	0	0	Normal
1	0	0	1	0	0	0	Normal
0	0	1	0	0	1	0	Normal
0	0	0	1	0	1	0	Normal
1	0	0	0	1	0	0	Drowsy
0	0	0	0	1	1	0	Drowsy
0	1	0	0	1	0	0	Danger
0	1	0	1	0	0	0	Danger
0	0	0	0	0	1	1	Danger

OE: Open Eye, CE: Closed Eye, T: Talking, CM: Closed Mouth, Y: Yawning, S: Sunglasess, HT: Head Tilt, DS: Drowsiness State

(a) (b) (c)

Figure 3 Decision matrix visualization (a) Drowsy (b) Normal (c) Danger

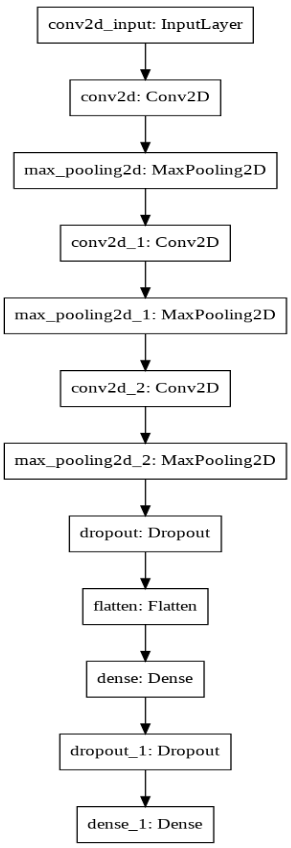


Figure 4 Driver drowsiness detection network

However, these values should be monitored for a certain period of time, as these instantaneous detections may lead to erroneous decisions. Therefore, this process is done for 50 frames. Then, these 50 lists are combined, and a decision matrix with 50×7 size is created. Finally, this decision matrix is resized to 50×50 for convolutional operations efficiency. Visualization of this decision matrix is shown in Figure 3.

The decision matrix is inferenced with a simple Convolutional Neural Network (CNN) as an image input. Proposed CNN is given in Figure 4. As shown in Figure 5, driver drowsiness detection is performed by processing every 50 overlapping frames from the start time. “Normal”, “Drowsy” and “Danger” classification decisions are obtained as a result of this approach.

4. EXPERIMENTAL RESULTS

The proposed method is trained with YawDD, NTHU and custom dataset. The object detection training loss curves are obtained for every 100 steps of the model and all training process is completed at 110000 steps (107 epoch). Training parameters used in the object detection method are given in Table 2. While training stage classification loss, localization loss and total loss are examined separately. These training loss curves are given in Figure 6.

The trained model is evaluated using COCO Object Detection Metrics. The mAP (mean Average Precision) is obtained as 0.949 which is calculated at 0.5 IoU (Intersection over Union). The comparison of the mAP values in this work with another method [10] which deals with the Driver Drowsiness Detection problem as an object detection is given in Table 3.

In driver drowsiness detection step, for training dataset, 202748 images of the normal class, 10226 of the drowsy class and 12160 of the danger class are collected from 383 videos. Training parameters used in the drowsiness detection method are given in Table 4. Test and training loss and accuracy curves are given in Figure 7. The hyperparameters in Table 2 and Table 4 are obtained using manual tuning approach. A small set of possible values of hyperparameters are determined. Then, these different sets of hyperparameters are manually performed and the set with best performance is chosen as final parameter set.

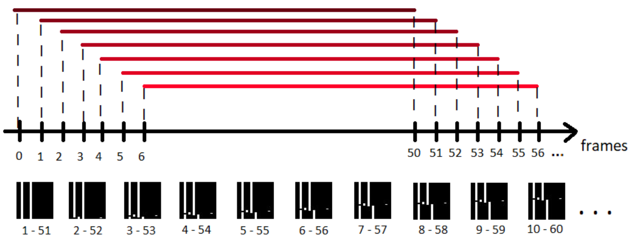


Figure 5 Flowchart of drowsiness detection processing

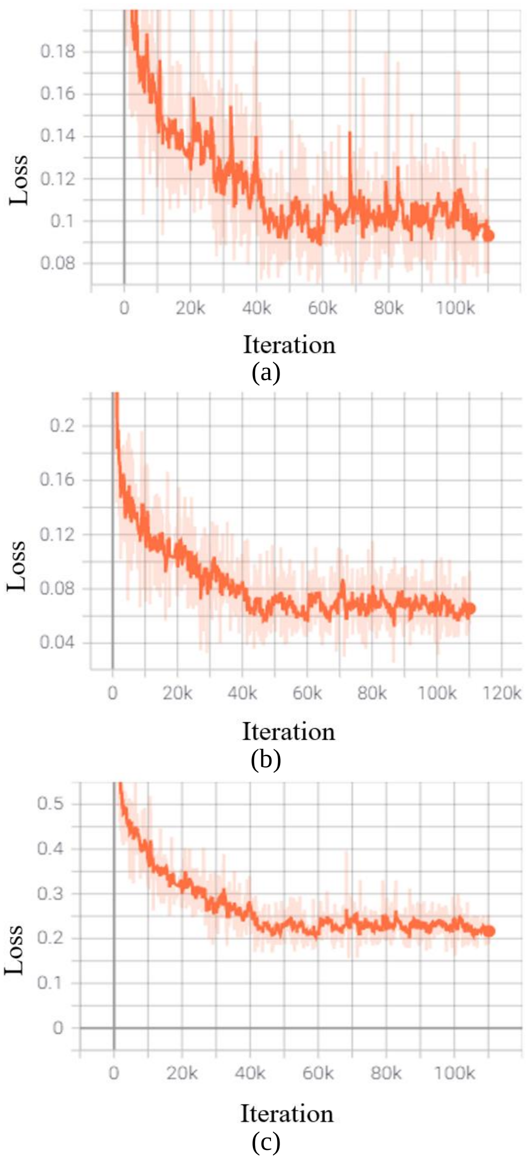


Figure 6 Training loss curves (a) Classification loss (b) Localization loss (c) Total loss

Table 2 Object detection network training parameters

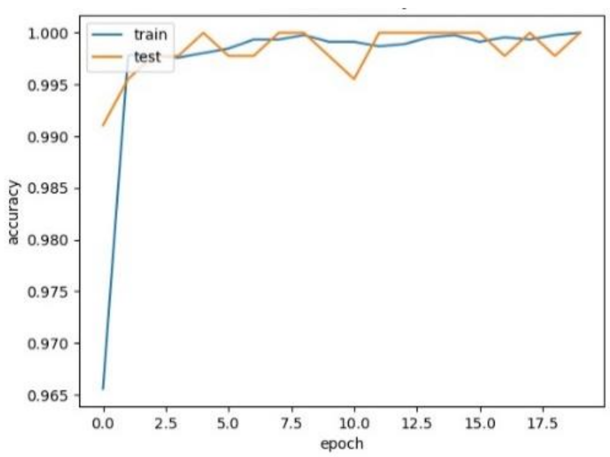
Parameter	Value
Learning rate	0.08
Epoch size	107
Batch size	6
Optimization algorithm	SGDM

Table3
Object detection evaluation on YawDD dataset

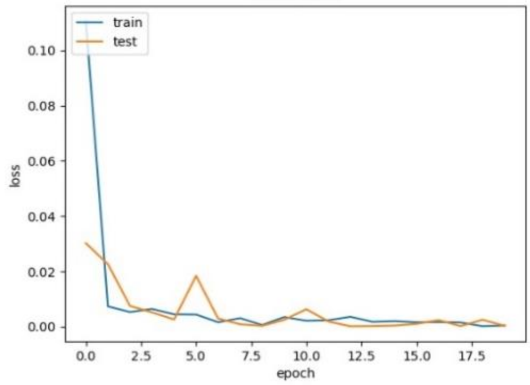
Method	mAP
[6]	0.837
Proposed Method	0.949

Proposed method is tested on custom dataset and NTHU-DDD dataset. Visual detection results on custom dataset are given in Figure 8. As seen from this figure, the proposed method is able to provide reliable detection results. Detailed quantitative evaluation on the NTHU-DDD dataset is given in Table 5.

Table 6 show objective evaluation in terms of accuracy of recent methods and proposed method on NTHU-DDD dataset. Jabbar et al. [8], Ramachandran et al. [9], Gaidar and Yakimov [10], Mehta et al. [5] and Jongmin Yu et al. [11] detect drowsiness situtation in the test conditions (with glasses, without glasses at night, with glasses at night, without glasses, with sunglasses) in NTHU-DDD dataset as same as proposed method. Accuracy results of these methods are directly taken from papers. As seen from Table 6, proposed method gives best result among the compared methods.

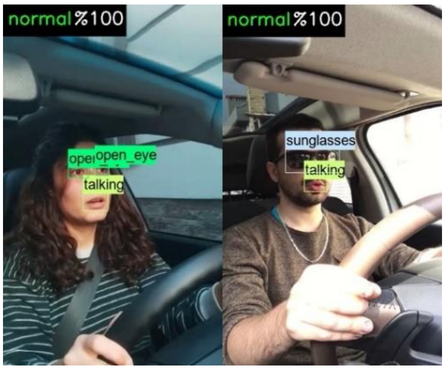


(a)



(b)

Figure 7 Training loss and accuracy curves (a) Accuracy curve (b) Loss curve



(a)

(b)

(c)

Figure 8 Visual driver drowsiness detection results (a) Normal (b) Drowsy (c) Danger

Table 4 Drowsiness detection network training parameters

Parameter	Value
Learning rate	0.01
Epoch size	20
Batch size	32
Optimization algorithm	SGDM

Table 5 Performance evaluation of proposed method on NTHU-DDD dataset

Category	Total frame number	Detected "drowsy" instead of "normal"	Detected "normal" instead of "drowsy"	Accuracy (%)
With glasses	38467	4518	275	87.19
Night time without glasses	39785	6637	480	82.39
Night with glasses	31141	5400	544	86.14
Without glasses	12450	741	320	92.12
With sunglasses	20319	4134	149	78.92
Total	142162	21430	1768	**85.35**

Table 6 Performance evaluation of proposed method and recent methods on NTHU-DDD dataset

Method	Accuracy (%)
Jongmin Yu et al. [11]	76.2
Ramachandran et al. [9]	79
Gaidar and Yakimov [10]	81
Jabbar et al. [8]	83.33
Mehta et al. [5]	84
Proposed Method	**85.35**

5. CONCLUSION

Most of the traffic accidents are caused by the carelessness of the drivers. An important part of these carelessness comes from drowsy driving. The aim of this work is to accurately determine the situations where the driver is normal, drowsy and in danger.

Mainly, driver drowsiness detection is handled as an object detection problem and drowsiness decision is provided by a simple Convolutional Neural Networks. In the proposed method, the fatigue decision is made as a result of an observation within a certain period of time, rather than instantaneous determinations. With this approach, erroneous detections are prevented by ensuring that every driver's eye closing or every opening his mouth is not perceived as drowsiness. In addition, one of the biggest problems of driver drowsiness detection, the inability to understand whether drivers with sunglasses are sleepy or tired, is solved by using the head position as an additional decision parameter. The experimental results show that the proposed drowsiness detection method is able to detect drowsy drivers in the challenging conditions with low computational complexity models which makes it suitable for real-time applications.

Funding

This work is supported by Kocaeli University Scientific Research Projects Coordination Unit under grant number 2019/021.

The Declaration of Conflict of Interest/ Common Interest

No conflict of interest or common interest has been declared by the authors.

Authors' Contribution

The first author contributed 40% and the second author 60% contributed to the study.

REFERENCES

[1] General Directorate of Police. URL http://trafik.gov.tr/kurumlar/trafik.gov.tr/04-Istatistik/Aylik/nisan21.pdf (accessed 5.22.21).

[2] Turkish Statistical Institute, URL https://data.tuik.gov.tr/Bulten/Index?p=Karayolu-Trafik-Kaza-Istatistikleri-2019-33628, (accessed 22.05.21).

[3] IIHSHLDI, URL https://www.iihs.org/topics/fatality-statistics/detail/state-by-state#fatal-crash-totals, (accessed 16.6.21).

[4] Cars, URL https://www.cars.com/articles/drowsy-driver-detection-systems-sense-when-you-need-a-break-1420684409199/#:~:text=BMW%20also%20allows%20you%20to,set%20it%20to%20%E2%80%9Cnormal.%E2%80%9D, (accessed 22.05.21)

[5] S. Mehta, S. Dadhich, S. Gumber, A. J. Bhatt, “Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio”, Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Jaipur - India, February 26-28, 2019.

[6] M.F. Shakeel, N.A. Bajwa, A.M. Anwaar, A. Sohail, A. Khan, Haroon-ur-Rashid, “Detecting Driver Drowsiness in Real Time Through Deep Learning Based Object Detection”, Advances in Computational Intelligence, IWANN 2019 Lecture Notes in Computer Science, vol 11506, 2019.

[7] R. Ahmed, K. E. Emon, M. F. Hossain, “Robust driver fatigue recognition using image processing”, 2014 International Conference on Informatics, Electronics & Vision (ICIEV), pp. 1-6, 2014.

[8] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, B. Hariri, YawDD, “A yawning detection dataset”, ACM International Conference on Multimedia Systems, Singapore, Singapore, March 19 - 21, 2014.

[9] C. H. Weng, Y. H. Lai, S. H. Lai, “Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network”, Asian

Conference on Computer Vision Springer, pp. 117-133, 2016.

[10] R. Jabbar, M. Shinoy, M. Kharbeche, K. Al-Khalifa, M. Krichen, K. Barkaoui, “Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application”, 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), pp. 237-242, 2020.

[11] S. Ramachandran, S. G. R. Rao, T. Sunder, “Drowsiness Detection Using Image Processing Techniques”, International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 6, 2018.

[12] A. I. Gaidar, P. Y. Yakimov, “Real-time fatigue features detection”, In: Journal of Physics: Conference Series, vol. 1368, no. 5, 2017.

[13] J. Yu, S. Park, S. Lee, M. Jeon, “Driver Drowsiness Detection Using Condition-Adaptive Representation Learning Framework”, IEEE Transactions on Intelligent Transportation Systems, vol. 20 no. 11, pp. 4206-4218, 2019.