SUST-DDD: A Real-Drive Dataset for Driver Drowsiness Detection

Esra Kavalcı Yılmaz
Sivas University of Science and Technology
Sivas, Turkey
esra.kavalci@sivas.edu.tr

M. Ali Akcayol
Gazi University
Ankara, Turkey
akcayol@gazi.edu.tr

***Abstract*—Driver drowsiness is one of the most important factors in traffic accidents. For this reason, systems should be developed to detect drowsiness early and to warn the driver by examining the driver or driving situations. This kind of systems play an important role to prevent traffic accidents. Three techniques are used to detect drowsiness: (1) based on vehicle parameters, (2) based on physiological parameters and (3) based on behavioral parameters. In this study, a new dataset for drowsiness has been created and some kind of deep learning methods such as AlexNet, LSTM, VGG16, VGG19, VGGFaceNet and hybrid deep networks have been applied on this dataset to predict drowsiness of the drivers. The experimental results show that the created dataset and implemented hybrid deep networks are successful to predict drowsiness with more than 90,53% for accuracy, 91,74% for precision, 91,28% for recall and 91,46% for f1score.**

I. Introduction

Every year, thousands of traffic accidents occur that cause financial losses, injuries and even deaths. In the 430,204 accidents that occurred in Turkey in 2021, 276,935 people were injured and 2,422 people died [1]. Traffic accidents are in the 8th rank among the causes of death in the world, and it is predicted that if no measures are taken, it may rise to the 5th rank [2]. According to the traffic statistics bulletin published by the Traffic Department of the General Directorate of Security in December 2021, 87% of the fatal-injury accidents in our country in 2021 are caused by the driver [1]. One of the most important factors of driver-related accidents is the driver's drowsiness.

Drowsiness is a physiological state of sleepiness. Among the most important symptoms of drowsiness are frequent blinking, eye rubbing, repetitive yawning, head tilt, distraction. Drowsiness causes loss of concentration and reduced reaction time. A driver in a state of drowsiness is unable to take action to prevent a possible accident [3]. In order to prevent such accidents, the development of various systems that recognize the signs of drowsiness, detect drowsiness and warn the driver plays an important role.

Techniques used for drowsiness detection are generally examined in three categories: (1) behavioral parameter-based techniques, (2) vehicle parameter-based techniques and (3) physiological parameter-based techniques. Behavioral parameters-based techniques are non-intrusive methods that measure drivers' drowsiness through driver behavioral parameters such as eye-closure rate, blinking, head position, facial expressions, and yawning. Vehicle parameters-based techniques are techniques that attempt to detect driver fatigue based on vehicle handling characteristics such as lane change frequency, vehicle speed variability, steering wheel angle, steering wheel grip strength. Physiological parameters-based techniques detect drowsiness based on drivers' physical conditions such as heart rate, pulse rate, respiratory rate, and body temperature. They are intrusive methods.

In this study, studies were carried out within the scope of methods based on behavioral parameters by using the new drowsiness detection dataset.Deep learning methods are used to predict the driver's drowsiness and the results of the study are compared.

This article is organized as follows. In Section 2, a literature review has been made about studies on driver drowsiness detection. In Section 3, driver drowsiness detection approaches are presented, and in Section 4, the datasets used for driver drowsiness detection and the dataset we presented in the study are presented comparatively. In Section 5, deep learning architectures used for video classification are examined. The details of the study are mentioned in Section 6. Finally, in Section 7, the conclusion are provide.

II. Related Work

Driver drowsiness is an important factor for both traffic and human safety. For this reason, there are some studies using different techniques and methods for the prediction of drowsiness.

Deng and Wu [4], proposed a system called DriCare that uses video images of the driver for driver drowsiness detection. The system basically addresses three critical challenges. First, because the driver's height will be different, the face positions in the video will also be different. In this case, it is important to monitor the position of the head. Second, they proposed a deep convolutional neural network to identify key points of the face, since the driver's eyes and mouth play an important role in detecting drowsiness. Third, they provided the determination of the driver's level of drowsiness. As a result of the study, it is seen that the best result is obtained in the light environment where the driver does not wear glasses.

Jabbar et al. [5], [6], stated that existing applications cannot be used in embedded systems due to limited computation and storage capacities, and they suggested models with smaller

ISSN 2305-7254

sizes. In their studies, they detected the landmarks on the driver's face from the images captured from the mobile device and transferred them to the deep learning model they developed. The models they developed are Drowsiness Detection based on Multi Layer Perceptron (MLP) and Facial Landmark Detection(D2MLP-FLD) model and Drowsiness Detection based on Convolutional Neural Network (CNN) and Facial Landmark Detection(D2CNN-FLD) model. As a result of the studies, they achieved 81% success in the D2MLP-FLD model and 83.33% in the D2CNN-FLD model, and stated that these developed models can be integrated into the dashboards of automobiles or mobile devices.

Chai et al. [7], in their study, determined the drowsiness of the driver according to the condition of the steering wheel. They evaluated 11 different parameters (steering angle, holding time of the steering wheel, angular velocity of the steering wheel, etc.) based on the parameters of the angle and angular velocity of the steering wheel. Afterwards, according to the results of the analysis of variance, five parameters (steering angle, standard deviation of steering angle, ratio of steering wheel holding time, steering wheel holding constant for more than 0.04 s, percentage of steering wheel angular velocity in the range of 2.5-5°/s) is selected. In the study, Multilevel Ordered Logit (MOL), Support Vector Machine (SVM) and Back Propagation (BP) Neural Network models were used.As a result of their studies, they obtained the highest success from the MOL model (72.92%) under the same classification conditions. The success values obtained are 63.86% for SVM and 62.10% for BP Neural Network.

Ko et al. [8], studied Electroencephalogram (EEG)-based driver drowsiness detection. In the study, they developed a deep convolutional neural network to detect drowsiness based on Differential Entropy (DE) obtained from EEG signals. In their study, they performed classification and regression experiments using the publicly available SEED-VIG dataset. With the VIGNet model developed as a result of the study, the accuracy value for the classification process was calculated as 96%, and the mean square error score (RMSE) for the regression process was calculated as 4%.

Sharma et al. [9], conducted a study of driver drowsiness based on EEG signals. In this training, they made an analysis based on Flexible Analytic Wavelet Transform (FAWT), which decomposes many frequency subbands. Feature extraction is done from the separated subbands. Then, the selection is made by statistical analysis. Different classification techniques were used to detect 'Drowsy' and 'Warning' states. As a result of the study, they obtained an accuracy of 95.6% with the model they proposed.

III. Drowsiness Detection Techniques

This section discusses techniques used for driver drowsiness detection. Three techniques are generally used for the detection of drowsiness.

A. Techniques Based on Behavioral Parameters

These are the techniques in which driver drowsiness is detected by using computer vision techniques after examining situations such as eyes open-closed state, blink rate, yawning and head movements through cameras where driver's facial expressions can be watched. Environmental factors such as lighting, brightness and road conditions are problems that will affect the accuracy and reliability of the measurement [10]. It is frequently preferred due to the absence of intrusive methods and low cost.

Dua et al. [11], proposed a new model consisting of four different CNN models (AlexNet, VGG-FaceNet, FlowImageNet, ResNet). The model aims to detect drowsiness using RGB videos of the drivers as input. These models take into account four different characteristics such as hand movements, facial expressions, behavioral features and head movements. AlexNet model for environmental changes such as night, day, indoor and outdoor; VGG- FaceNet model to extract facial features such as gender and ethnicity; FlowImageNet for behavioral features and head movements; ResNet is used for hand gestures. The National Tsuing Hua University Drowsiness Detection Dataset (NTHU-DDD) was used in the study and 85% success was achieved as a result of the study.

Yazdi and Soryani [12], investigated the yawning state of the driver to detect drowsiness in their study. Their work consists of three phases. First of all, using the depth information, the minimum depth is accepted as the nose, and the picture is divided so that the lower half of the face will be used in the following stages. In the second stage, the mouth area in the lower half of the face and whether the mouth is open or closed are determined. At this stage, the detection of an open or closed mouth achieved 86% accuracy. Finally, the driver's drowsiness is determined by comparing the height of the open mouth with a threshold value. In the study conducted on the data containing 50 yawns and 40 half-open or closed mouth images, the yawn of the individual was determined with an accuracy rate of 95%. Compared to other methods, it has been stated that this algorithm is advantageous in that it is not sensitive to day and night lighting changes. However, in this study, if the minimum depth point changes as a result of the driver's head moving, the performance will decrease.

B. Techniques Based on Vehicle Parameters

These are the techniques in which driver drowsiness is detected by examining vehicle driving characteristics such as lane change frequency, steering wheel movements, vehicle speed change using sensors placed on the steering wheel and accelerator/brake pedals. Steering wheel movement and standard deviation of lane position are the most frequently preferred parameters [13]. These systems are costly as they require special equipment. In addition, it is easily affected by external factors such as weather and road conditions. For these reasons, it is not preferred much.

Arefnezhad et al. [14], presented a system based on deep neural networks that detects driver drowsiness using vehicle-based parameters. In the study, they evaluated 5 parameters:

lateral deviation from the road center line, lateral acceleration, yaw rate, steering angle and steering wheel speed. The study was carried out in a fixed-base driving simulator (Graz Advanced Driving Simulator, ADSG) at the Graz University of Technology (TU Graz). The study was carried out in a fixed-base driving simulator (Graz Advanced Driving Simulator, ADSG) at the Graz University of Technology (TU Graz). As a result of the study, the best result was obtained from the CNN-LSTM model with an accuracy rate of 96%.

C. Techniques Based on Physiological Parameters

These are intrusive techniques in which drowsiness is detected by monitoring physiological signals such as EEG signals, ECG (Electrocardiogram) signals, HRV (Heart Rate Variability) of the driver with various measuring devices placed on the driver. These techniques are more reliable as the biological characteristics of the driver are evaluated. However, the devices placed on the driver disturb the driver, so they are not preferred.

Fujiwara et al. [15], propose a driver drowsiness detection algorithm based on HRV analysis and validate the proposed method by comparing it with EEG-based sleep scoring. The performance of the proposed algorithm was evaluated using a driving simulator. Experimental data were collected from 34 participants, 25 male and 9 female, while driving a virtual vehicle on a simulator. HRV data obtained at the end of the study and sleep onset were evaluated according to EEG data determined by a sleep specialist. As a result of the evaluation, it was observed that drowsiness was detected in 12 of the 13 sleep levels specified as transitional sleep or light sleep before sleep onset.

Yaacob et al. [16], stated that EEG is important in measuring drowsiness levels because it shows the electrical activity of the brain, and they used EGG signals to analyze driver drowsiness. In the study, firstly, feature extraction was performed by filtering the alpha frequency (frequency band between 8-13 Hz) from the EEG signals. Then, the extracted features were classified using the decision tree. As a result of the study, they achieved success in the range of 77.1%-97.20%.

IV. Datasets for Drowsiness

There are many data sets to be used in driver drowsiness detection studies based on behavioral parameters. A few of the frequently used datasets that are publicly available are briefly described below. In TABLE I, the data set properties are briefly stated.

- **ULg Multimodality Drowsiness Database (DROZY):** A total of 14 people, 3 men and 11 women, participated in data collection. Care was taken to ensure that people were not addicted to alcohol, using drugs or having sleep disorders. A sleep diary was kept for the participants to verify their sleep needs. Experiments were carried out at regular intervals in a quiet, isolated laboratory environment. During this process, attention was paid to avoid warning substances such as tea and coffee. In addition to the video recordings of the participants in the data set, there are four different electrical biosignal values: electroencephalogram (EEG), electrooculography (EOG), electrocardiogram (EKG) and electromyogram (EMG) [17].
- **The NTH Drowsiness Detection Dataset (NTHU-DDD):** The dataset was collected by the National Tsuing Hua University Computer Vision Laboratory. The dataset consists of videos of 36 people. Pertaining to drowsiness such as yawning, nodding, slow blinking; Action images not related to drowsiness, such as laughing, talking, looking at both sides, were recorded. Images were taken under day and night lighting conditions. Images of participants were recorded while sitting in a chair and playing a simple driving game with simulated steering wheel and pedals. While recording the videos, they were instructed by someone else to make a series of facial expressions [18].
- **The Real-Life Drowsiness Dataset (RLDD):** The dataset was obtained from a total of 180 videos, each video being approximately 10 minutes, in 3 different states of each of 60 participants as alert, low vigilant and drowsy. Participants consisted of 51 men and 9 women. Participants were given 20 days to prepare three videos and were asked to record videos in the environment they were in (at home, work, school, etc.) when they felt alert, low vigilant or drowsy. The participants were asked to take the videos from an arm's length away, in a way that the face could be seen and in accordance with the driving video. The videos were taken by the participants via their mobile phones or webcams. In addition to the advantages of the data set, it has a significant disadvantage that it does not include the driving moment [19].
- **Yawning Detection Dataset (YAWDD):** The dataset was created in two different sets, from the camera placed on the front mirror of the car and the camera placed on the driver's panel. The dataset consists of videos of 57 male and 50 female participants in three different situations: normal driving (no talking), talking or singing while driving and yawning while driving in a parked vehicle. It is a suitable data set for face and mouth recognition studies, as well as being useful for yawn detection [20].
- **National Cheng Kung University Driver Drowsiness Dataset (NCKUDD):** The dataset consists of videos of 25 participants taken during normal driving with a camera placed in front of the driver. The videos were shot both in daylight and in dark environments and include normal, sleepy, distracted, talking while driving, eating while driving, talking on the phone while driving, and other abnormal driving behaviors.While the fact that the dataset includes real driving moments can be considered as an advantage, it includes situations that may endanger driving safety, such as talking on the phone or eating while driving [21].
- **Sivas University of Science and Technology Driver Drowsiness Dataset (SUST-DDD):** When the existing data sets are examined, it is seen that the data are

generally collected in simulation, laboratory, cafe, school etc. environments. These datasets include videos of participants pretending to be driving or sitting around doing nothing. In some data sets consisting of real driving moments, it is seen that there are situations that may endanger the safety of the driver and the traffic, such as using cell phone, eating something, etc. while driving. The presented SUST-DDD, on the other hand, is a dataset consisting of videos recorded by the cameras of their phones when drivers feel tired and normal during real driving. During the videos, the drivers recorded a video at any moment of the road they were going. The route used is not a fixed route determined by the authors and varies for each participant. Participants were not asked to engage in any particular behavior while driving, thus ensuring a completely realistic and safe driving experience.

19 participants, aged between 24 and 42, were asked to take a video with their own phones placed in front of the driver's seat when they felt drowsy/normal. In this case, since the drivers do not have to leave their comfort zone, their reactions/behaviors are more realistic. As for shooting hours, the hours when daylight is low, such as coming to work or leaving work, were preferred. After the videos with different lighting conditions, different sizes and resolutions recorded by each participant with their own phone were collected from the participants, the data set was created by dividing the videos into pieces, each 10 seconds and 224x224x3.

Finally, the 2074 videos obtained were watched by the jury consisting of 3 volunteer participants and were labeled as drowsy-not drowsy by voting. The distribution of video numbers in the dataset is shown in TABLE II.

TABLE I. PROPERTIES OF DATASETS

Dataset	**Number of Participant**	***Place***	***Situation***
DROZY	14	Lab	Normal Drowsy
NTHU-DDD	36	Game Simulation	Yawning/Nodding Slow Blinking Laughing/Talking Looking at Both Sides
RLDD	60	Home/Work /School etc	Alert Low Vigilant Drowsy
YAWDD	107	Parked Vehicle	Normal Talking or Singing Yawning
NCKUDD	25	Driving Instant Vehicle	Normal/Sleepy Distracted/Talking Eating/Drowsy Using Cell Phone Other
SUST-DDD	19	Driving Instant Vehicle	Drowsy Not Drowsy

TABLE II. NUMBER OF SUST-DDD VIDEOS

Data	**Drowsy**	**Not Drowsy**	**Total**
Video	975	1099	2074

V. VIDEO CLASSIFICATION WITH DEEP LEARNING

Deep learning is based on computational systems that imitate the human brain. Thanks to deep learning, audio, text and image data are scanned and interpreted by the computer model. Deep learning models are used for various tasks such as classifying, detecting, and locating objects.

Liu stated that video classification based on deep learning should be strengthened when compared to image processing. Based on this idea, he examined the video classification processes based on deep learning in detail in his studies [22].

CNNs are structures with powerful feature extractor. Recurrent Neural Network (RNN) structures, on the other hand, are structures that use sequential information and draw conclusions by comparing the information previously introduced to the system with the new information. Long Short Term Memory (LSTM), which is often used with time series data, is a special type of RNN that can learn long term dependencies. In many video classification studies [23], [24], [25], feature extraction is done primarily with CNN architectures. Afterwards, the results obtained are processed with the temporal information from RNN and classification processes are performed.

VI. THE EXPERIMENTAL RESULTS

In this section, the performance of the SUST-DDD dataset is evaluated using different architectures. A dataset consisting of videos shot with different cameras, including different genders, ages, backgrounds, was collected.

Due to the different sizes of the videos collected from the participants, the sizes of the videos were initially set to 224x224x3. Afterwards, 20 frames were obtained from each 224x224x3 video and preprocessed. Training was carried out using deep learning models such as AlexNet, VGG16, VGG19 and VGGFaceNet, with 25% of the data set for test and 75% for training. Relu activation function is used in all of the CNN models where the training is carried out. In max pooling layers, pool size is used as 3x3 for AlexNet, while pool size is set to 2x2 for other models. There are 2 fully connected layers in VGG16, VGG19 and AlexNet, and 3 in VGGFaceNet. The outputs obtained as a result of the training were fed into the LSTM architecture and classification processes were carried out. The LSTM models consists of an LSTM layer, 2 fully connected layers, and a softmax. Relu and Sigmoid activation functions were used as activation functions, respectively. The detail of the LSTM structure is shared in Fig.1. During the classification processes, k-fold cross validation was used to avoid biases and errors that may arise from the distribution of the data set. In the k-fold cross validation process, k=4 was set.

Layer (Type)	Output Shape	Parameters
LSTM	(None, 512)	9439232
Dense	(None, 512)	262656
Activation	(None, 512)	0
Dense	(None, 64)	32832
Activation	(None, 64)	0
Dense	(None, 2)	130
Activation	(None, 2)	0

Fig. 1. The Structure of LSTM

TABLE III. RESULTS OF THE MODELS

Model	Accuracy	Precision	Recall	F1 Score
VGG19+LSTM	%90,53	%91,74	%91,28	%91,50
VGG16+LSTM	%89,39	%91,81	%89,09	%90,42
AlexNet+LSTM	%63,91	%63,78	%97,91	%77,24
VGGFaceNet+LSTM	%84,94	%83,65	%94,92	%88,92

As a result of the study, it is seen that the most successful model in prediction of drowsiness is the VGG19+LSTM hybrid model with 90.53% for accuracy, 91.74% for precision, 91.28% for recall and 91.46% f11score values. The most unsuccessful results were obtained with the AlexNet+LSTM hybrid model with 63.91% for accuracy, 63.78% for precision, 97.91% for recall and 77.24% f11score values. Results for all models used are shown in TABLE III. The graph where the results of all models are shared is shown in Fig.2.

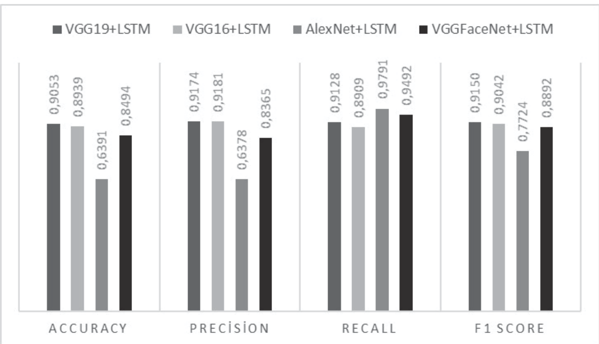

Fig. 2. The Graphical Display of Results

VII. CONCLUSIONS

In this paper, firstly, a driver drowsiness dataset consisting of real driving moments is presented. Compared to other commonly used driver fatigue datasets, the presented dataset appears to be a more realistic dataset containing safer driving moments. During the creation of the dataset, it was aimed to increase the reliability of the dataset by performing the classification processes in the presence of the jury. In the second stage of the study, the performance of the dataset was evaluated by using various deep learning models. After training the presented dataset with deep learning architectures such as AlexNet, VGG16, VGG19 and VGGFaceNet, classification processes were performed with LSTM architecture. As a result of the study, it is seen that the most successful result was obtained with the VGG19+LSTM model as 90.53%.

ACKNOWLEDGMENT

The experiments reported in this study were carried out with the workstation of Sivas University of Science and Technology.

We would like to thank the participants who volunteered for dataset collection and the jury members who volunteered for the classification procedures.

REFERENCES

[1] Traffic Department of the General Directorate of Security, Available at: http://trafik.gov.tr/istatistikler37 .
[2] M. Özen and F. Zorlu, "An analysis of crash rates and the crash pattern on state highways in Turkey," Tek. Dergi/Technical J. Turkish Chamb. Civ. Eng., vol. 29, no. 5, pp. 8589–8604, Sep. 2018, doi: 10.18400/tekderg.308318.
[3] E. E. Galarza, F. D. Egas, F. M. Silva, P. M. Velasco, and E. D. Galarza, "Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone," Adv. Intell. Syst. Comput., vol. 721, no. April, pp. 563–572, 2018.
[4] W. Deng and R. Wu, "Real-Time Driver-Drowsiness Detection System Using Facial Features," IEEE Access, vol. 7, pp. 118727–118738, 2019, doi: 10.1109/ACCESS.2019.2936663.
[5] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, "Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques," Procedia Comput. Sci., vol. 130, pp. 400–407, 2018, doi: 10.1016/j.procs.2018.04.060.
[6] R. Jabbar, M. Shinoy, and M. Kharbeche, "Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application," pp. 237–242, 2020.
[7] C. Meng, L. Shi-Wu, S. Wen-Cai, G. Meng-Zhu, and H. Meng-Yuan, "Drowsiness monitoring based on steering wheel status," vol. 66, no. July 2018, pp. 95–103, 2019, doi: 10.1016/j.trd.2018.07.007.
[8] W. Ko, K. Oh, E. Jeon, and H. Suk, "VIGNet: A Deep Convolutional Neural Network for EEG-based Driver Vigilance Estimation," 2020 8th Int. Winter Conf. Brain-Computer Interface, no. February, pp. 1–3, 2020, doi: 10.1109/BCI48061.2020.9061668.
[9] S. Sharma, S. K. Khare, V. Bajaj, and I. Ahmad, "Improving the separability of drowsiness and alert EEG signals using analytic form of wavelet transform," Appl. Acoust., vol. 181, p. 108164, 2021, doi: 10.1016/j.apacoust.2021.108164.
[10] M. Ramzan et al., "A Survey on State-of-the-Art Drowsiness Detection Techniques," vol. 7, 2019.
[11] M. Dua, Shakshi, R. Singla, S. Raj, A. Jangra, "Deep CNN models-based ensemble approach to driver drowsiness detection," Neural Comput. Appl., vol. 33, no. 8, pp. 3155–3168, 2021, doi: 10.1007/s00521-020-05209-7.
[12] M. Z. Jafari Yazdi and M. Soryani, "Driver Drowsiness Detection by Yawn Identification Based on Depth Information and Active Contour Model," 2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019, pp. 1522–1526, 2019, doi: 10.1109/ICICICT46008.2019.8993385.
[13] M. K. Hussein, T. M. Salman,A. H. Miry, M. A. Subhi, "Driver Drowsiness Detection Techniques: A Survey," vol. 2021, no. Bicits, 2021.
[14] S. Arefnezhad et al., "Expert Systems with Applications Applying deep neural networks for multi-level classification of driver drowsiness using Vehicle-based measures," Expert Syst. Appl., vol. 162, p. 113778, 2020, doi: 10.1016/j.eswa.2020.113778.
[15] K. Fujiwara et al., "Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG," IEEE Trans. Biomed. Eng., vol. 66, no. 6, pp. 1769–1778, 2019, doi: 10.1109/TBME.2018.2879346.

[16] S. Yaacob, N. I. Z. Muhamad'Arif, P. Krishnan, A. Rasyadan, M. Yaakop and F. Mohamed, "Early driver drowsiness detection using electroencephalography signals," pp. 2020–2023, 2020.

[17] Q. Massoz, T. Langohr, C. François and J. G. Verly, "The ULg multimodality drowsiness database (called DROZY) and examples of use," 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1-7, doi: 10.1109/WACV.2016.7477715.

[18] C. Weng, Y. Lai, S. Lai, "Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network", In Asian Conference on Computer Vision Workshop on Driver Drowsiness Detection from Video, Taipei, Taiwan, Nov. 2016

[19] R. Ghoddoosian, M. Galib, and V. Athitsos, "A realistic dataset and baseline temporal model for early drowsiness detection," IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., vol. 2019-June, pp. 178–187, 2019, doi: 10.1109/CVPRW.2019.00027.

[20] Abtahi S, Omidyeganeh M, Shirmohammadi S, Hariri B. YawDD: A yawning detection dataset. InProceedings of the 5th ACM multimedia systems conference 2014 Mar 19 (pp. 24-28).

[21] C. Chiou et al., "Driver Monitoring Using Sparse Representation With Part-Based Temporal Face Descriptors," vol. 21, no. 1, pp. 346–361, 2020.

[22] M. Liu, "Video Classification Technology Based on Deep Learning," 2020 International Conference on Information Science, Parallel and Distributed Systems (ISPDS), 2020, pp. 154-157, doi: 10.1109/ISPDS51347.2020.00039.

[23] M. A. Russo, L. Kurnianggoro, and K. H. Jo, "Classification of sports videos with combination of deep learning models and transfer learning," 2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE 2019, pp. 7–9, 2019, doi: 10.1109/ECACE.2019.8679371.

[24] Y. Yang, R. A. Sarkis, R. El Atrache, T. Loddenkemper, and C. Meisel, "Video-Based Detection of Generalized Tonic-Clonic Seizures Using Deep Learning," IEEE J. Biomed. Heal. Informatics, vol. 25, no. 8, pp. 2997–3008, 2021, doi: 10.1109/JBHI.2021.3049649.

[25] M. Kopaczka et al., "Assessment of Laboratory Mouse Activity in Video Recordings Using Deep Learning Methods," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 3673-3676, doi: 10.1109/EMBC.2019.8857807.