A Review of the Computer Engineering Literature Relating to Drowsy Driver Detection System

Volkan Mazlum, Mert Alp Kuvandık,
Bahu Tongal

Department of Computer Engineering, University of Cankaya

Advisor: Serdar ARSLAN

Nov 10, 2022, version 0.1

Abstract

The Drowsy Driver Detection System is a project that aims to avoid deadly traffic accidents. Driver sleepiness is one of the leading causes of traffic accidents. They get weary or exhausted and may fall asleep because of long-distance travel, stress, or sleep deprivation. We want to avoid and decrease such accidents by developing a sleepy detecting system. There are many projects developed on the Drowsy Driver Detection System and dozens of articles. Because drowsiness and fatigue are some of the main problems that cause accidents. Finding solutions to these both reduces the pain burden on the health systems of countries and can significantly reduce the number of deaths in accidents. Different algorithms and solutions developed for these have been published in the articles I just mentioned. These algorithms mainly focus on extracting certain features from facial expressions such as yawning, eye, mouth opening, and head movements. The study closes with a comprehensive bibliography that includes brief descriptions of each publication. Our main goal is to focus on trying to prevent accidents via the methods we will use and to be used by us and focus on reducing the number of deaths with new algorithms of our system.

1 Introduction

The Drowsy Driver Detection System is a program that controls and processes the drivers' states and movements, and facial expressions and draws the road accordingly. What we mean by drawing a road is that it includes different reactions according to the driver's situation. The system itself must decide this and keep the driver awake. On the other hand, it is thought to be a system that can give warnings to stop the driver from driving and to get enough sleep. In general, it is a project that every driver candidate will need and even should have.

According to the World Health Organization (WHO), millions of people die or become disabled in traffic every year in the world. According to statistics, the main cause of these accidents is driver fatigue and carelessness. According to the report of the American Automobile Association, 7% of all accidents and 21% of fatal accidents are due to driver fatigue. According to another study conducted in the USA, 2% of the accidents that occurred in 2009 were caused by users being sleep-deprived. Also, according to a 2017 Organization for Traffic Safety survey, 42.4% of drivers drove without at least one or more days of sleep, fewer than six hours of sleep, and severe for 87.9% of drivers. [1]

Sleeping during a trip is not something that happens suddenly. Specifies a process. During this process, some helleri indicate different situations such as yawning, the head falling forward, and the eyes being closed to closing. These need to be determined beforehand. Our work will also focus here. Our main goal is to examine the eye and mouth states with different formulas and algorithms and to detect the sleep state early. If this is achieved, many unfortunate situations are likely to be avoided.

In Section 2, we offered an overview of our research on the Drowsy Driver Detection approach and the problems that we can run into. In Section 3, we discussed our overall impressions of the literature, with a focus on the basic algorithms and different methods that we discovered. The following section also includes a condensed list of "Top Articles" that offer a detailed overview of the papers we evaluated, and we thought that they are very appropriate for our research. In Section 5, we organized our collection of papers further into the challenge areas that each article tries to solve. We conclude with an annotated bibliography that captures the most recent relevant research in the field of Computer engineering.

2 Drowsy Driver Detection

Using machine learning/deep learning algorithms, there are numerous studies on drowsy driving detection systems. The two types of monitoring systems that make up the real-time detection system are those that are the driver- and vehicle-oriented. Using information gathered by the car's sensors, such as the location of the vehicle in the lane, steering wheel motions, pedal pressure, and other factors, the car-oriented device analyzes the driver's habits to determine whether the driver is sleepy or not. The driver-oriented system uses facial traits to detect sleepiness. The human face is animated and capable of displaying a variety of emotions. Being one of the most noticeable features of the human face, human eyes are crucial for object recognition and facial expression studies. It is helpful to be able to identify eyes when identifying facial features. Recognizing eyes is helpful when identifying features of a face. The data used in the articles is not the same for every article. For example, in the paper in [7], the authors used facial expressions to develop the model. In the article in [9], eye ratio and facial expressions are used as input. In the article in [16], since the relationship between steering and the level of sleepiness is tried to be examined, steering data is used as input. In the article in [8], the authors tried to find and predict the distraction level and drowsiness level using smart glasses. As we mentioned above, although human eyes, faces, and similarities are of critical importance, different data are used to develop models.

We are considering using computer vision algorithms used for face recognition. In addition, using artificial neural networks (CNN), we will ensure that the project includes deep learning algorithms and learn the status of the driver from the data itself. We will check the accuracy of this learning with different metrics and try to maximize it. In addition, we will make it permanent through an android or web application by using the models derived because of this. Thus, the system will become usable for a long time. Although I said only CNN above, different methods are used in different articles that we will see in the future. To give a short example of these, RNN and RF algorithms are used in the article [11]. The articles in [9] and [10] used CNN, the method we are talking. In the article in [13], unlike the others, ANN combined with DFT was used. The articles in [18] used artificial neural networks (ANN), support vector machines (SVM), and adaptive boosting to improve accuracy of model. As a final example, the article in [7] can be given. QDNN (Quantum Dot Neural Network) algorithm was used to achieve the desired results. In short, different algorithms can be used according to the purpose of the article. Figure-1 can be shown in the below that you can examine results of articles.

Article	Accuracy	Algorithms
Real time Driver Detection [1]	98.81%	M-ConNN
Automatic Driver Detection [13]	72.25%	ANN–EAR–MAR
Driven Facial Emotion Recognition [7]	99.31%	QDNN
Real-time Driver Drowsiness [2]	95.2% - 93.5%	CNN- MTCNN
Driver Drowsiness Detection [9]	83%	CNN
Eye Detection for a Real-time [18]	-	ANN–SVM–AdaB
Driver Safety Detection [3]	99%	FD–NN–TL-VGG
Drowsiness Detection Techniques [4]	-	VSM–HMM–SVM

Figure-1 Algorithms are used in articles

2.1 Challenges of Drowsy Driver Detection System

Some problems may occur during the implementation phase of this project. We need to anticipate them. Some of the problems that may occur in the project are:

- Errors that may occur when using the pre-defined dataset
- Difficulty finding training data
- Errors that may occur while performing Face Recognition

If these problems occur, they will be tried to be overcome with different engineering methods and the project will be successful.

2.1.1 Finding Appropriate Dataset and Pre-defined Dataset Problems

This problem has been mentioned in most articles used in the literature. Every dataset published so far has issues or deficiencies, albeit different from each other. For instance, no videos were taken in the dark in one, or a data set was not created for individuals using glasses in the other. Due to these problems, when the system encounters such data, the accuracy value will decrease, and its reliability will decrease. To avoid such issues, both the datasets in the literature are used, and a dataset created by us is intended to be used in our project. Thus, we will bring a new dataset to the literature and try to increase the evaluation values of the project.

2.1.2 Face Recognition Problem

While using the obtained datasets, the videos will be framed by different libraries used for computer vision such as OpenCV. It is aimed to make face detection on the obtained frames. However, the fact that the person in the photo has glasses or the photos are taken in dark environments can easily prevent detection. To prevent these, different algorithms or different formulas that can be found in the literature that can detect eye and mouth deficits can be used. These formulas will be applied to the system by using the facial landmarks to be obtained and calculations will be made.

2.2 Field of This Project and Algorithms that are used

The fields of this project are mainly Software and its engineering and Artificial

Intelligence. According to the methods to be used in its content, computer vision, Deep Neural Networks (CNN, LSTM, etc.) can be added. This list of fields can be further developed with different algorithms to be added or added. In addition, this project will be developed in python language and basic python libraries such as Scikit-learn, Keras, Tensorflow, PyTorch, and YOLO will be used. Some of the methods and libraries mentioned here need to be explained and detailed. In another part, we will examine different methods and libraries.

2.2.1 CNN

For picture identification, CNN, or the Convolutional Neural Network object detection technique, is utilized. In this approach, we choose distinct areas from various portions of a picture and attempt to categorize the presence of an item in that region. The CNN algorithm, on the other hand, can only select a restricted number of areas.

2.2.2 YOLO

The YOLO (You Only Look Once) algorithm is quite different from prior approaches. It has been taught to do bounding box prediction and classification simultaneously. As a result, the YOLO algorithm is quicker than the CNN method. However, YOLO has similarities with earlier algorithms. It also has certain disadvantages. Because each grid may only have two bounding boxes, YOLO may fail if items are too close together. Furthermore, because to spatial constraints, it has difficulty identifying tiny objects.

2.2.3 Scikit-learn, Keras and Tensorflow

Scikit-learn is a popular Python tool for data research and machine learning. It supports a wide range of operations and offers a variety of algorithms. Scikit-learn also includes documentation for its classes, methods, and functions, as well as the algorithms that are employed.

TensorFlow is a deep learning math library that is open source. TensorFlow, which is part of the machine learning cluster, also supports classical machine learning methods and was created largely for internal Google usage by the Google Brain team. TensorFlow is defined as a deep learning (Deep learning) library with open-source code. With a single API (Application Programming Interface), TensorFlow's flexible structure allows calculations on all systems.

Keras is a deep learning package that works atop the previously stated TensorFlow. It is a Python package that does not define nearly any deep learning model. It is a high-level neural network API that can be used with not just TensorFlow but also Theano and CNTK.

2.2.4 RNN - LSTM

The results of people's experiences and activities provide some opportunity for learning. To move on, these inferences are stored in memory and reused as necessary. In neural networks, fresh outputs are produced and stored in memory while some inputs are used repeatedly as needed. Traditional neural networks lack a system for remembering previously learned information. Recurrent neural networks (RNNs) are intended to analyze input in time and in sequence and to transfer information in various steps. A neural network can analyze data that is presented as a time series and has a relationship between the individual time steps.

A typical step inside the RNN is differentiated by the LSTM. This difference is that new parameters are created and stored via a variety of mathematical processes, as opposed to a conventional basic neural network construction. So, when necessary, it is possible to obtain the input or output data from the preceding steps. By transmitting the data from the previous layer to the necessary ports and performing mathematical operations, it simultaneously stores an output value and delivers it as an input to the next layer.

3 Main findings

While scanning articles, most of the articles prepared on the drowsy driver detection system focused on 3 basic features. There are 3 basic measurement techniques used to develop the Drowsy driver detection system. These are Image measure, biological measure, and vehicle measure. First, let's examine the image measure techniques, which we will also use.

3.1 Image Measure on Drowsy Driver Detection

It is aimed to conclude by focusing on the face and eye features through photography. They try to examine the driver's drowsiness with different algorithms (like CNN-LSTM). In addition, factors such as blinking and yawning are also used, apart from the state of the plain and face and the state of the eyes.

3.2 Biological Measure on Drowsy Driver Detection

Psychological-Biological scales that can be obtained by methods such as EEG, and the other is physical scales that focus on facial, mouth, and body movements. Some models give very good results in both. These sibyls are one of the methods used to further strengthen the system for drowsy driver detection systems. Examples of these signals are factors such as brain activity, heart rate, breathing rate, and body temperature. It can produce more accurate results than those used in the Image method. Examples of these methods are EEG, ECG, and PPG. In short, to be clear, EEG examines and categorizes brain signals. ECG examines the electrical activity on the skin. PPG, on the other hand, focuses on factors such as cam pressure and oxygen content. Again, algorithms such as LSTM can be used with these data.

3.3 Vehicle Based Measure on Drowsy Driver Detection

These measurements focus on the analysis and study of driving patterns. That is, situations such as tire movements are examined. But it is not very preferred to be used alone. It is generally used in conjunction with other measurement methods. Examples of these measurements are the steering wheel and the lateral distance.

As we can understand from the articles in the literature, the main problem in this project is the diversity of the datasets. Some datasets only contain images in bright environments, while others have recordings in dark environments. Some of them also have different labels such as using and not using glasses. Therefore, the datasets used also significantly affect the results of the article. In short, data set selection is of great importance. Generally, NTHU Drowsy Driver Detection (NTHU-DDD) Dataset has been used in the literature. For example, this dataset is used in the article in [5]. Also, in the article in [1], YawdDD and NTHU are used together, because those datasets are a large content dataset. In [15], [16] articles, different datasets are used for different purposes. After a general explanation, we move on to the part where we will examine in detail some of the articles that we consider very important and selected from the general literature review. Figure-1 can be shown in the below that you can examine.

Article	Dataset is used
Real time Driver Detection [1]	YawdDD and NTHU

Automatic Driver Detection [13]	Their own datasets
Driven Facial Emotion Recognition [7]	FER-2013 and CK+
Driver Drowsiness Detection [5]	NTHU
Driver Safety Detection [3]	Mixture dataset of ZJU

Figure-2 Dataset is used in articles

4 Top papers

Although each article has valuable information, we selected some of the "top papers" from our literature review. Our aim is to collect the existing data in a general framework. In each article, different problems were encountered, and different solutions were sought. Examination of these guides us in our own project. Therefore, it is important to re-examine each article step by step. But in order not to waste time describing the whole literature, we will examine a few in detail to provide general information.

Our first article was written by Wanghua Dengi and Ruoxue Wu [2], and the DriCare system, which can detect fatigue status states such as yawning and blinking from video recordings, is proposed. In addition, unlike the algorithms used in other literature, a new face-tracking algorithm has been developed to improve accuracy. This model has 68 basis points. They argue that by using these points, the driver's status can be easily determined. They say that an accuracy of 92% is obtained from the model developed using features obtained from the mouth and eye. The model basically detects eye closure, blinking, and yawning. In addition, the Multiple Convolutional Neural Networks (CNN)-KCF (MC-KCF) algorithm was used to obtain better results in low-light situations. Multitask convolutional neural networks (MTCNN) are used to detect some situations that the KCF algorithm cannot detect. Another method used is CNN. They used CNN to detect the state of the eyes. In short, human face tracking is done with the help of MC-KCF from the data obtained from the video stream. Then, feature extraction is performed, including mouth and eyes. As the last stage, there is the evaluation part.

In the article [3], they tried to make a real-time application using CNN. Accuracy has been tried to be increased and the algorithm has been tried to be accelerated as much as possible. Also, a Fully Designed Neural Network (FD-NN), Transfer Learning in VGG16 and VGG19 with extra designed layers (TL-VGG) were used to make eye status classification. Considering the scarcity of the dataset on eye closure detection in the literature, a different dataset (mixture dataset of ZJU) is presented in the article. It has been used as 3 category drowsiness detectors as Vehicle-based, Signal-based and Facial feature-based. In the decision-making phase, unlike other articles, they adjusted the system according to the actual psychological values. For example, for a healthy person, blinking is between 100-400 ms. If this number increases to 1 second, it is perceived by the system as a symptom of fatigue. In addition, if the person is detected blindfolded in 12 consecutive photographs in the data set, the warning system works. Finally, a very high value for accuracy, such as 99%, is achieved.

In another article [4], the methods found in the literature are divided into 3 categories: Behavioral, Vehicular, and Physiological parameters-based techniques. Apart from these, the most used supervised learning algorithms in the literature were tried to detect drowsiness. PERCLOS was used for eye-based drowsiness detection. In the article, SVM is used for mouth and yawning classification. The main purpose of using SVM is to reduce the cost. For different situations such as poor lighting, the driver wearing classes, and the Driver with a mustache, experiments were made according to each categories mentioned above. The most appropriate category for each situation was found, providing a basis for further research. Although mainly focused on SVM, CNN and HMM are also used. The differences between them are revealed, so HMM shows a less error rate, but both CNN and HMM are slow in training and expensive as compared to the SVM classifiers. The main purpose of this article is to get results by trying different algorithms and categories for different situations. These results were considered by the authors to form a basis for further research.

In the article [5], deep architecture called deep drowsiness detection is proposed. This structure includes 3 deep networks to obtain environmental variations and learn local facial movements and head gestures. The results of these 3 networks are combined with softmax. An accuracy of around 73.06% was achieved. The networks used are AlexNet, VGG-FaceNet, and FlowImageNet. It was used to collect features related to AlexNet drowsiness. VGG-FaceNet has been used to obtain facial features for different situations such as ethnicities, and hairstyles. FlowImageNet was used to find out whether facial and head movements from sequential images were related to drowsiness. Every three networks are independently fine-tuned for multi-class drowsiness classification. In addition, 2 different ensemble strategies were used. One of them is independently averaged architecture (IAA), and the other is feature-fused architecture (FFA). The NTH Drowsy Driver Detection dataset is used in the article.

In another article [1], studies were carried out on 2 different datasets. These are the YawdDD and NthuDDD datasets. Unlike other articles, it is recommended by the authors to use Multi-tasking Convolutional Neural Network (ConNN*) to find drowsiness in this article. As a feature of this algorithm, both eye and mouth information can be categorized at the same time. This feature provides the model with both great speed and storage space. Calculations of the duration and percentage of closed eyes (PERCLOS) as well as the frequency and duration of mouth and yawning sneezes are used to assess driver fatigue (FOM). As a result of this study, a very high value of 98.81% was obtained. The proposed system can simulate how the interaction between the mouth, eyes, and face.

This article [11] explains the technology developed to detect drowsy behavior based on the result that drowsy driver exhibits more drowsy behavior. Since it is difficult to measure physiological signals in the developed behavior detection technology, it is preferred to use driving information. Driving information was obtained from experiments using simulators. The input parameters used consist of lateral acceleration, longitudinal acceleration, steering angles. These parameters were investigated in different time periods. ANN and RF were used as algorithms. It turned out that the RF algorithm is better at examining data sets than the ANN algorithm. The studies conducted in the article provided a high accuracy of 84.8% and it was concluded that it can be used in real vehicles.

In the article [17], These characteristics are included in the new sleepy driving prevention system that is proposed in the current development. The suggested system, known as a wakefulness-keeping support system (WKSS), is made up of an active gaming system (AGS), which corresponds to the second and third characteristics, and a drowsiness detection system (DDS), which corresponds to the first feature. These characteristics are included in the new sleepy driving prevention system that is proposed in the current development. The suggested system, known as a wakefulness-keeping support system (WKSS), is made up of an active gaming system (AGS), which corresponds to the second and third characteristics, and a drowsiness detection system (DDS), which corresponds to the first feature.

According to the research [12], 23.5% of traffic accidents were caused by the driver being sleepy. This article focuses on two models to find a solution to this problem. Two models are proposed in this research. The first model can determine whether a driver is sleepy. The first model is successful in detecting whether the driver is drowsy, while the second Model is designed to detect if the driver yawns, blinks or speaks. A total of 12 different situations were detected and defined by the second model. Indicators showing that the driver is sleepy are defined using multitask learning. The accuracy of the model was 92.40%. The second model is a multi-class model that can recognize the driver's condition, such as whether they are chatting, blinking, or yawning. Driver drowsiness signs are discovered through multitasking learning.

In the last article [6], a different parameter was used compared to other articles. Gold standard brain bio-physiological signal and facial expression data were used to detect early drowsiness. Data were obtained from a group of 10 people. ANN was used to process them and reveal a model. Different features such as EAR, MAR, and FL are used. (ie, eye aspect ratio (EAR), mouth aspect ratio (MAR), face length (FL)) It has been suggested that the model created with ANN together with these features gives the best results.

These articles are part of our overall literature review. You can find many more detailed articles in the References section.

5 Conclusion

Drowsy Driver Detection System has recently become the most important vehicle part used by vehicle companies. Because the number of people who lost their lives in traffic is also increasing. Therefore, R&D resources transferred to these systems are increasing day by day and research are increasing. Therefore, every effort and time given to this project is very valuable for us. As we can see from our literature review, ANN and its derivative algorithms are generally used in this field. Because they both give very good results and are less expensive than other algorithms. In addition, facial expressions are mainly used for drowsiness detection, although there are approaches in more than one category. We will focus on this in our project. Another issue is the dataset problem. Although there is more than one dataset in the market, we have found the most suitable one thanks to this research. With different evaluation results, the best dataset became easily selectable. Finally, this literature review we have done has been written to form a general idea and to form a basis for the project to be done.

References

[1] B. K. Savaş and Y. Becerikli, "Real Time Driver Fatigue Detection System Based on Multi-Task ConNN," in *IEEE Access*, vol. 8, pp. 12491-12498, 2020

[2] Deng, W., & Wu, R. (2019). Real-time driver-drowsiness detection system using facial features. Ieee Access, 7, 118727-118738.

[3] Hashemi, M., Mirrashid, A., & Beheshti Shirazi, A. (2020). Driver safety development: Real-time driver drowsiness detection system based on convolutional neural network. SN Computer Science, 1(5), 1-10.

[4] Ramzan, M., Khan, H. U., Awan, S. M., Ismail, A., Ilyas, M., & Mahmood, A. (2019). A survey on state-of-the-art drowsiness detection techniques. IEEE Access, 7, 61904-61919.

[5] Park, S., Pan, F., Kang, S., & Yoo, C. D. (2016, November). Driver drowsiness detection system based on feature representation learning using various deep networks. In Asian Conference on Computer Vision (pp. 154-164). Springer, Cham

[6] Inkeaw, Xu, L., Li, S., Bian, K., Zhao, T., & Yan, W. (2014, February). Sober-Drive: A smartphone-assisted drowsy driving detection system. In 2014 International conference on computing, networking, and communications (ICNC) (pp. 398-402). IEEE.

[7] A. Mustafa Hilal, A., Elkamchouchi, D. H., Alotaibi, S. S., Maray, M., Othman, M., Abdelmageed, A. A., ... & Eldesouki, M. I. (2022). Manta Ray Foraging Optimization with Transfer Learning Driven Facial Emotion Recognition. Sustainability, 14(21), 14308.

[8] -J. Chang, L. -B. Chen and Y. -Z. Chiou, "Design and Implementation of a Drowsiness-Fatigue-Detection System Based on Wearable Smart Glasses to Increase Road Safety," in *IEEE Transactions on Consumer Electronics*, vol. 64, no. 4, pp. 461-469, Nov. 2018, doi: 10.1109/TCE.2018.2872162.

[9] 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," *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 2020, pp. 237-242, doi: 10.1109/ICIoT48696.2020.9089484.

[10] BenJi, Y., Wang, S., Zhao, Y., Wei, J., & Lu, Y. (2019). Fatigue state detection based on multi-index fusion and state recognition network. IEEE Access, 7, 64136-64147.

[11] Wang, M. S., Jeong, N. T., Kim, K. S., Choi, S. B., Yang, S. M., You, S. H., ... & Suh, M. W. (2016). Drowsy behavior detection based on driving information. International journal of automotive technology, 17(1), 165-173.

[12] Seetharaman, R., Sridhar, S., & Mootha, S. (2020, October). Detection and State Analysis of Drowsiness using Multitask Learning with Neural Networks. In 2020 Fourth International Conference on Intelligent Computing in Data Sciences (ICDS) (pp. 1-8). IEEE.

[13] Inkeaw, P., Srikummoon, P., Chaijaruwanich, J., Traisathit, P., Awiphan, S., Inchai, J., ... & Theerakittikul, T. (2022). Automatic driver drowsiness detection using artificial neural network based on visual facial descriptors: pilot study. Nature and science of sleep, 14, 1641-1649.

[14] Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: a review. Sensors, 12(12), 16937-16953.

[15] Dwivedi, K., Biswaranjan, K., & Sethi, A. (2014, February). Drowsy driver detection using representation learning. In 2014 IEEE international advance computing conference (IACC) (pp. 995-999). IEEE.

[16] Eskandarian, A., & Mortazavi, A. (2007, June). Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection. In 2007 IEEE intelligent vehicles symposium (pp. 553-

559). IEEE.

[17]Ibe, T., Fujiwara, K., Hiraoka, T., Abe, E., & Yamakawa, T. (2020). Development of Game-Like System Using Active Behavior Input for Wakefulness-Keeping Support in Driving. IEEE Transactions on Intelligent Vehicles, 6(2), 323-332.

[18]Coetzer, R. C., & Hancke, G. P. (2011, June). Eye detection for a real-time vehicle driver fatigue monitoring system. In 2011 IEEE Intelligent Vehicles Symposium (IV) (pp. 66-71). IEEE.