JJEE
Volume 6, Number 3, 2020
Pages 205-215
Jordan Journal of Electrical Engineering
ISSN (Print): 2409-9600, ISSN (Online): 2409-9619

# Camera-Based Driver Monitoring System for Abnormal Behavior Detection

**Ahmad Aljaafreh***

Communications, Electronics and Computer Engineering Department, Tafila Technical University, Tafila, Jordan
E-mail: a.aljaafreh@ttu.edu.jo

*Received: April 08, 2020* *Revised: April 30, 2020* *Accepted: May 14, 2020*

***Abstract—*** Psychological and physiological status has a big impact on the driver's behavior. It affects the driver's visual scanning behavior which helps drivers maintain visual attention. This paper proposes a system for detecting the abnormal driving behavior from the sequential pattern of the driver's peripheral visual scanning. The system continuously monitors the driver's activities through an in-vehicle camera to measure the driver's visual distraction. Feature descriptors of both the transition and rotation vectors of the driver's head pose and eye gaze are extracted and provided to a linear support vector machine (SVM) classifier to output one of six driver's common gaze zones. Then, a reservoir computing (RC) based on echo state networks (ESNs) is used for driver behavior classification from the sequence of the driver's gaze zones. The system is implemented on NVIDIA Jetson Nano to execute the processing of all the data since it has a Maxwell graphics processing unit (GPU) with 128 compute unified device architecture (CUDA) cores. The obtained results show that the driver's behavior can be classified to normal or abnormal based on his visual scanning activities with high accuracy. They also demonstrate the efficiency of both SVM and ESN in detecting the abnormal driver's behavior from a sequence of driver's gaze zones. Moreover, the results show that the proposed monitoring system is capable of detecting the driver's abnormal behavior with a detection accuracy of 98%, making it an appropriate candidate for successful deployment in both driver assistant systems (DAS) and driving safety support systems (DSSS).

***Keywords—*** Monitoring system; Behavior detection; Driver behavior; Classification; Support vector machine; Echo state network; Driver assistant system; Driving safety support system.

## 1. INTRODUCTION

Every year about 1 to 1.24 million people are killed and 20 to 50 millions are injured on roads' accidents across the world due to abnormal driving behaviors and poor road environments [1, 2]. Driving behavior is a notion, related to how the driver runs the car in the traffic scene and nearby environment [3]. Driver's visual scanning behavior is a smart practice to view the total traffic scene. Good behavior helps drivers to maintain attention and to be prepared for any dangers or traffic congestions. Driver's psychological and physiological circumstances affect driving behavior and could lead to driver's distraction due to driving for long-distance, drowsiness, fatigue or cluttered environments [4, 5].

The expansion of road networks along with the increase of road accidents have led to a need for driver assistant systems (DAS) and driver safety support systems (DSSS) [5]. These systems are typically equipped with sensors and digital cameras and installed inside vehicles to support, monitor and alert drivers. Computer vision and artificial intelligence techniques can work with collected data of both road environment and driver behavior to ease driving and increase driver's safety [6].

Driver's behavior in this context can be measured by different parameters. Head pose and gaze estimation of the driver are key parameters for predicting driver's behavior and thus, alerting him appropriately without distraction [7]. The gaze dynamics are correlated to

* Corresponding author
Article's DOI: 10.5455/jjee.204-1586348076

the driver's activities. Gaze dynamics are shrunk to a pattern of durations and frequencies based on driver's activities.

## 2. BACKGROUND AND SIGNIFICANCE

Driver's behavior is reviewed in this section based on gaze estimation and higher semantic. The pattern of the driver's gaze is studied in [8] through maneuvering in a freeway driving. The paper explored lane change for left and right as well as lane-keeping.

A camera is installed inside the car to continuously monitor driver's activities and the driver's non-front-facing frames were identified from videos using the Viola-Jones algorithm [9]. Inattention is recognized by repeated having non-frontal face consequent frames. Horizontal mean intensity of the eye region is used to recognize drowsiness through monitoring eye status. The support vector machine (SVM) classifier was used in [10] to identify the driver gaze direction from the location and scales of face parts.

A large labeled dataset is used in [11] to answer the question: can the head pose data without gaze data be used to predict the gaze region? Authors in [12] take a step towards the generalized system that is invariant to different subjects, perspective, and scale using convolutional neural network (CNN). Checking of the mirror is analyzed in [13] under different driving and maneuver conditions. Results showed that mirror checking can be considered as a good indicator in recognizing driver behavior. In [14], a near-infrared camera along with a deep learning gaze detection model is used without initial user calibration. The model was evaluated on their database and the open Columbia gaze dataset (CAVE-DB). A system is proposed for gaze region detection in [15] without gaze calibration. The system combined head pose and multi-scale eye images. A framework that combines both head and eye information is proposed in [16] to estimate the gaze region in both day time and night time.

A system for continuous driver gaze zone estimation in real-world driving conditions is proposed in [17]. It combines both multi-zone iterative closet points (ICP) based head pose tracking and appearance-based gaze estimation. In [18], a monitoring system is proposed to study the correlation between the driver's pose direction and the road scene.

This paper assumes a correlation between the driver abnormal behavior and the peripheral visual scanning sequential pattern. Computer vision and artificial intelligence are employed to understand driver attention. A fixed in-vehicle camera captures videos of the driver to track the driver's head pose and gaze. This work proposes a system to detect the abnormal driving behavior from the sequential pattern of the peripheral visual scanning. Feature descriptors of both the transition and rotation vector of the head pose and gaze are provided to a classifier to output one of the gaze zones. The sequence of gaze zones is supplied to the echo state network (ESN) model to detect abnormal behavior. ESN is a recurrent neural network with a sparsely connected hidden layer (usually with connectivity of 1 %). The connectivity of the hidden neurons and their weights are fixed and allocated randomly. The weights of the output neurons can be learned so that the network can generate different temporal patterns or replicate them.

 

The main contributions of this paper are summarized as:

- The classification of driver's gaze zones using linear SVM which classifies head pose and gaze vectors to one of the six common gaze regions.
- The classification of driving behavior to normal/abnormal using ESN for gaze zone sequence.
- Building the alert system rules.

The rest of this paper is structured as follows: section 3 presents a description of the system, employed to detect abnormal driving behavior. Section 4 presents results and discussion of the classification techniques and section 5 closes with conclusions.

## 3. SYSTEM DESCRIPTION

The proposed monitoring system detects the abnormal driving behavior from the sequential pattern of the driver's peripheral visual scanning. It works in five main stages: a) video capturing of the driver; b) estimation of the driver's head pose; c) estimation of the driver's pupil location; d) classification of the head pose and pupil location to one of the six common gaze zones (road, right mirror, left mirror, rear mirror, dashboard, and center console); and finally e) classification of the sequence of gaze zones to normal or abnormal driving behavior. The block cycle of the aforementioned five consequent stages of the system is shown in Fig. 1.

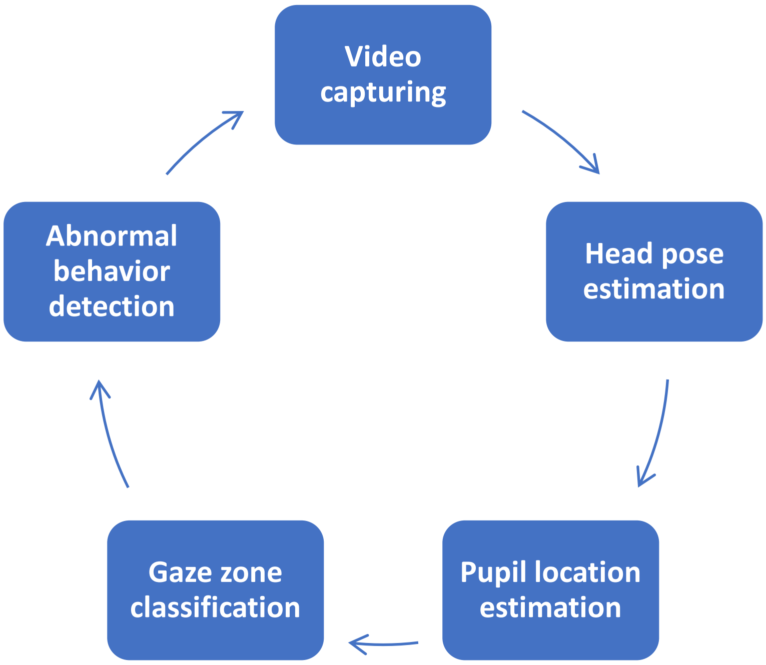


Fig. 1. Block cycle of the five consequent stages for the driver's monitoring system.

### 3.1. Video Capturing

The video is captured by a Raspberry Pi camera module. The used module is interfaced with the NVIDIA Jetson Nano minicomputer board to execute and process the captured image frames since it has a Maxwell GPU with 128 CUDA cores.

### 3.2. Driver Head Pose Estimation

The driver's head pose is the relative orientation and position of the driver's head to the camera. Head pose is represented by 6 parameters that are divided equally to three for translation and three for rotation. An in-vehicle fixed camera is directed toward the driver's face to be used for head pose estimation. The head pose has been estimated as follows:

a) A face detector was implemented to detect the face of the driver. The detected face image is extended for step two.
b) A trained deep learning model based on tensor flow was adopted for 68 facial landmark detection.
c) Perspective-n-point (PNP) algorithms are used to calculate the head pose based on the facial landmark.
d) Head pose data is filtered by a Kalman filter to smooth the data.

### 3.3. Gaze Estimation

Gaze estimation aims to find the location of the iris's center, which is the center of the black area inside the eye. The eye image is extracted using the eyes' enclosing landmarks. Iris is detected by applying a pixel intensity threshold after converting the eye image to grayscale. Then, the center coordinates are computed as the gaze estimation.

### 3.4. Gaze Zone Classification

The input of this stage is a feature descriptor of the transition and rotation vector of the head pose in addition to the location of the pupil. The feature descriptor is provided to SVM classifier to output one of the common six gaze zones which are road, right mirror, left mirror, rear mirror, dashboard and center console as depicted in Fig. 2. A linear SVM - with linear kernel function, automatic kernel scale, and a one-vs-one multiclass method- has been used. The model is cross-validated to prevent the overfitting of data.

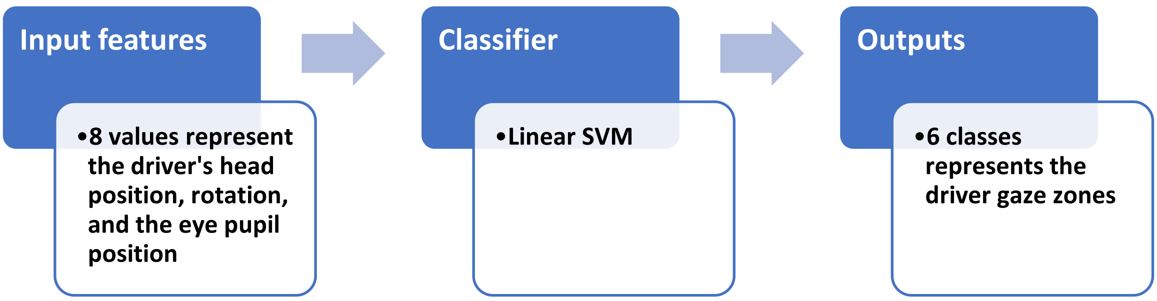


Fig. 2. Workflow of the driver's gaze zone classification.

An SVM is a representation of the input data as space points, mapped in a way that enables distinguishing it - by as wide as possible simple distance – to different categories. SVM classify data by class labels given a set of examples by finding a hyperplane which separate the data point of each class. SVM choose the hyperplane that maximizes the distance between the hyperplane and the closest points in each feature space region which are called support vectors. So, the unique optimal hyperplane is the plane that maximize this distance.

 

### 3.5. Driver's Abnormal Behavior Detection

The driver's abnormal behavior detection is modeled as a sequence classification problem. The sequence of the gaze zones is supplied to ESN recurrent neural network model to detect abnormal behavior as shown in Fig. 3. The figure shows the schematic illustration of the proposed driver's behavior detection workflow including the sequence classification stage. ESN recurrent neural network is trained to classify the sequence of the gaze zones to either normal or abnormal classes. The training data set was generated by simulation. ESN is trained with 4000 sequences of normal and abnormal driving divided as 2000 sequences, each of length 200 samples, for each driving behavior. The performance of the ESN model has been evaluated with three different datasets.

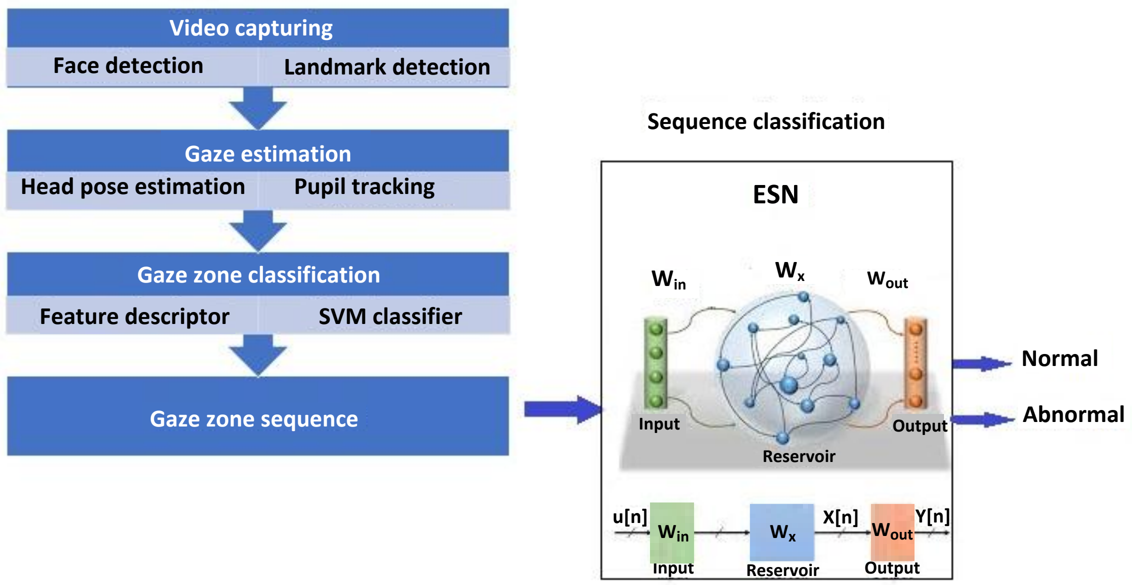


Fig. 3. Workflow of the driver's behavior classification.

#### *3.5.1. Gaze Zones Sequence Generation*

The sequence of the gaze zones is generated by simulation for both normal and abnormal driver's behavior. The assumptions of the probabilities for each gaze zone are demonstrated in Table 1. The probability of the road zone is the highest since the driver spends most of his time looking to the road. The probabilities for the dashboard and center console are the least.

Table 1. The assumptions of the probabilities for each driver's gaze zone.

| Gaze zone | Road | Rear mirror | Left mirror | Right mirror | Dashboard | Center console |
|---|---|---|---|---|---|---|
| Probability | 0.6 | 0.1 | 0.1 | 0.1 | 0.05 | 0.05 |

The only two differences between the normal and abnormal gaze zone sequences are:

a) For normal driving: the time difference between two consequent road gaze zones is not greater than two seconds.

b) For normal driving: the time difference between two consequent same mirror checking is not greater than eight seconds.

Sequences of length 200000 are generated for both normal and abnormal behavior. Then, these sequences are used to train the ESN recurrent neural network.

*3.5.2. Gaze Zones Sequence Classification*

Recurrent neural networks using echo state networks (ESNs, A.K.A reservoir computing) are used for the gaze zone sequence classification. ESN is a recurrent neural network with a loosely connected hidden layer, called a 'reservoir'. ESN works amazingly fine with the existence of random time series. ESN is used in research to speeds up the training process of the neural networks [19]. ESN is fast and takes minutes to train the network with a data set of 2000 sequences where each sequence has 200 samples.

## 3.6. Alert System

The system is capable of giving alerts to the drivers according to the rules, shown in Fig. 4, as follows:

a) If the system detects abnormal driving behavior for some time.
b) If the time difference between two inconsequent roads gaze zones is greater than two seconds. A glance of two seconds away from the road is enough to distract the driver's attention to unsafe conditions [20, 21].
c) If the time difference between two inconsequent rear view and side-view mirrors checking is greater than eight seconds [22].

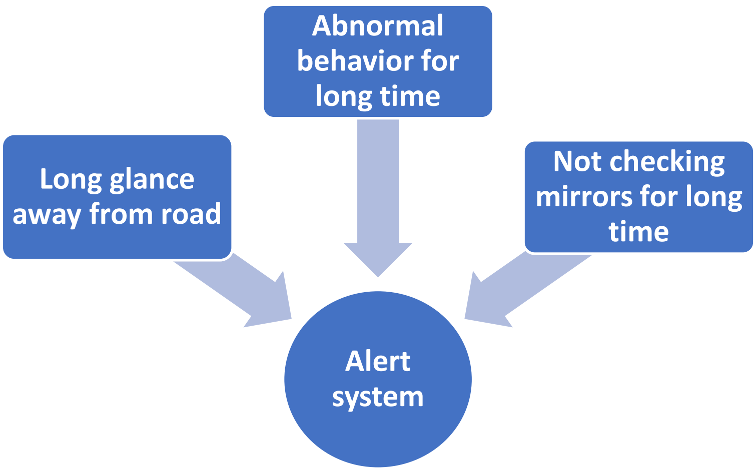


Fig. 4. Alert system rules.

## 4. RESULTS AND DISCUSSION

In this section, both stages of the gaze zone classification and the driver behavior classification are described. The driver's behavior is classified into one of two classes, namely normal and abnormal. The proposed monitoring system has been tested on simulated training data. Both gaze zone classification and driver behavior classification models have been built on training data and tested on another set of data.

 

### 4.1. Gaze Zone Classification

The classification was achieved based on the feature descriptor to provide and select an output among the common six gaze zones. Five different classifiers were evaluated based on the classification accuracy as in Table 2. These classifiers are Linear SVM, Coarse KNN, Gaussian Naive Bays, Linear Discriminant and Coarse Decision Tree. A linear SVM with linear kernel function, automatic kernel scale, and a one-vs-one multiclass method gave the best accuracy. It was trained for 15 s with 11377 of labeled data. The model was cross-validated with five folds to protect the model from the overfitting of the data. The classification accuracy was 99.2% as in the confusion matrix in Fig. 5. It is clear that the confusion is in the dashboard gaze region, rear mirror region, center console region, and the road region since these regions are so closed to each other.

Table 2. Classification accuracy for five different classifiers.

| Classifier name | Classification accuracy [%] |
|---|---|
| Linear SVM | 99.2 |
| Coarse KNN | 98.7 |
| Gaussian Naive Bays | 97.3 |
| Linear Discriminant | 95.0 |
| Coarse Decision Tree | 67.7 |

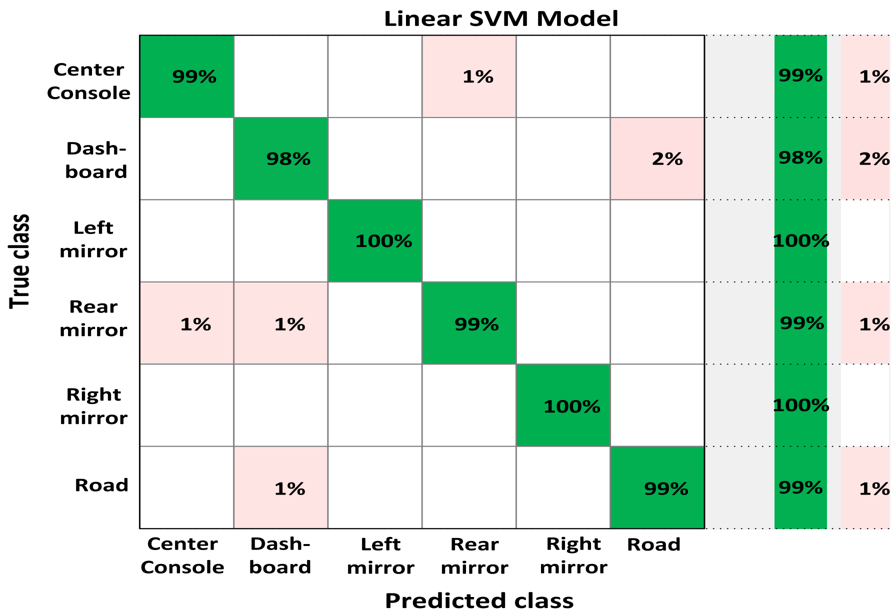


Fig. 5. The confusion matrix of the linear SVM classifier.

### 4.2. Driver Behavior Classification

Two different performance metrics - represented in [23] - are used to evaluate the results of this investigation. The first is the accuracy (ACC) which is the most commonly used classification performance metric. It is the ratio between the correctly classified samples to the total number of the samples. ACC is defined as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative.

The second performance metric is the error rate (ERR). It represents the misclassified numbers from both classes. It is the complement of the classification accuracy as demonstrated in Eq. (2);

$$\mathrm{ERR} = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \quad (2)$$

The ESN model has been evaluated as shoen in Table 3. The classification ACC and the ERR are shown in the table for the training data and for two newly generated input data named as input 1 and input 2. The input training data set achieved 98.3% classification ACC, whereas the newly generated data achieved 98% and 98.25% classification ACC, respectively.

ESN recurrent neural network is good enough to detect the abnormal driver behavior from the sequence of the driver's gaze zones. ESN classifier performance is described in Figs. 6-8. The confusion matrices show the number of true positive, true negative, false positive and the false negative. ESN model is tested on three different datasets. Each dataset has 4000 sequences. The first one is used for the training process and testing as well. The other two are generated just for testing. Firstly, ESN classified 1950 of the sequence as TP and classified 1981 as TN for the training data as in Fig. 6. Secondly, it also classified 1941 of the gaze zones sequences as TP and classified 1979 as TN for the newly first generated sequences (input 1) as in Fig. 7. Lastly, it classified 1949 of the input sequences as TP and classified 1981 as TN for the newly second generated sequences (input 2) as in Fig. 8.

Table 3. Performance analysis of ESN classifier.

| Sequence (2000X200) | ACC [%] | ERR [%] |
|---|---|---|
| Input (training) | 0.983 | 0.017 |
| New input 1 | 0.980 | 0.02 |
| New input 2 | 0.9825 | 0.0175 |

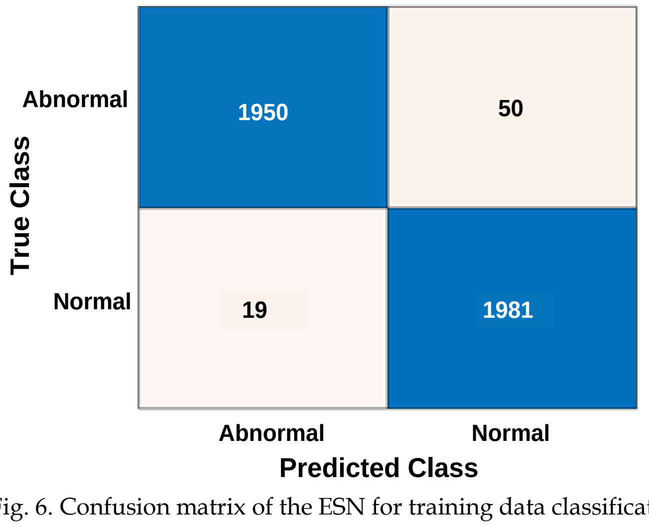


Fig. 6. Confusion matrix of the ESN for training data classification.

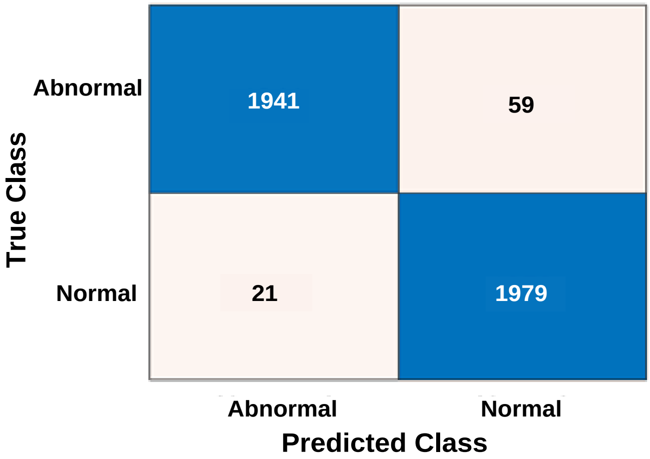


Fig. 7. Confusion matrix of the ESN for input 1 data classification.

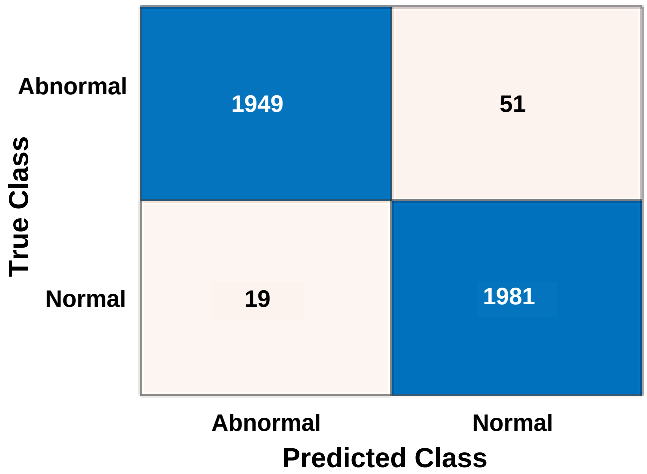


Fig. 8. Confusion matrix of the ESN for input 2 data classification.

## 5. CONCLUSIONS

In this work, we proposed a monitoring system of the driver's behavior using modern classification techniques. The proposed system was applied to data, obtained from an in-vehicle digital camera. It can be deployed in both DAS and DSSS to help, support and alert drivers on possible abnormal behavior situations during the driving task. Detection of the driver's abnormal behavior was achieved by analyzing the temporal gaze patterns of the driver. The proposed system consists of five consecutive stages where different detection and classification techniques were deployed in each one.

The results of our experiments showed that the proposed method can expect driver's abnormal behavior with an accuracy of 98%. Both Linear SVM and ESN classifiers achieved excellent performance results in the gaze zone classification and sequence classification, respectively.

Future work aims to design and build a suitable embedded advanced driver assistance system (ADAS) that can alert drivers when abnormal behavior is detected. Also, other metrics of driver's behavior can be addressed considering the mental concentration of the driver. Other features can also be added to evaluate the road environment around the vehicle.

## REFERENCES


[1] M. Regan, J. Lee, K. Young, *Driver Distraction: Theory, Effects, and Mitigation,* CRC Press: Boca Raton, FL, USA, 2008.

[2] M. Khan, S. Lee, "A comprehensive survey of driving monitoring and assistance systems," *Sensors*, vol. 19, no. 11, pp. 2574, 2019.

[3] Z. Elassad, H. Mousannif, H. Al Moatassime, A. Karkouch, "The application of machine learning techniques for driving behavior analysis: a conceptual framework and a systematic literature review," *Engineering Applications of Artificial Intelligence*, vol. 87, pp. 103312, 2020.

[4] J. Abukhait, "Image-based visibility estimation of road signs in cluttered environment," *10th International Conference on Computer Engineering and Applications*, Barcelona, Spain, pp. 221-226, 2016.

[5] J. Abukhait, "Visibility estimation of road signs considering detect-ability factors for driver assistance systems," *WSEAS Transactions on Signal Processing*, vol. 12, no. 1, pp. 111-117, 2014.

[6] J. Abukhait, I. Abdel-Qader, J. Oh, O. Abudayyeh, "Road sign detection and shape recognition invariant to sign defects," in *2012 IEEE International Conference on Electro/Information Technology*, Indianapolis, IN, pp. 1-6, 2012.

[7] K. Doman, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, H. Murase, Y. Tamatsu, "Estimation of traffic sign visibility toward smart driver assistance," *2010 IEEE Intelligent Vehicles Symposium*, San Diego, CA, pp. 45-50, 2010.

[8] S. Martin, S. Vora, K. Yuen, M. Trivedi, "Dynamics of driver's gaze: explorations in behavior modeling and maneuver prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no .2, pp. 141-150, 2018.

[9] A. Modak, S. Paradkar, S. Manwatkar, A. Madane, A. Deshpande, "Human head pose and eye state based driver distraction monitoring system," *Proceedings of 3rd International Conference on Computer Vision and Image Processing*, pp. 403-421, Singapore, 2020.

[10] M. Chuang, R. Bala, E. Bernal, P. Paul, A. Burry, "Estimating gaze direction of vehicle drivers using a smartphone camera," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, OH, pp. 165-170, 2014.

[11] L. Fridman, P. Langhans, J. Lee, B. Reimer, "Driver gaze region estimation without use of eye movement," *IEEE Intelligent Systems*, vol. 31, no .3, pp. 49-56, 2016.

[12] S. Vora, A. Rangesh, M. Trivedi, "On generalizing driver gaze zone estimation using convolutional neural networks," *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, pp. 849-854, 2017.

[13] N. Li, C. Busso. "Detecting drivers' mirror-checking actions and its application to maneuver and secondary task recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no .4, pp. 980-992, 2015.

[14] R. Naqvi, M. Arsalan, G. Batchuluun, H. Yoon, K. Park, "Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor," *Sensors*, vol. 18, no. 2, pp. 456, 2018.

[15] G. Yuan, Y. Wang, T. Zhao, X. Ding, Z. Mi, X. Fu, "Eye gaze region estimation via multi-scale sparse dictionary learning," *14th IEEE Conference on Industrial Electronics and Applications*, Xi'an, China, pp. 1459-1463, 2019.

[16] C. Luo, *Driver's Gaze Zone Estimation in Realistic Driving Environment by Kinect*, Doctoral Dissertation, University of Ottawa, 2018.

[17] Y. Wang, G. Yuan, Z. Mi, J. Peng, X. Ding, Z. Liang, X. Fu, "Continuous driver's gaze zone estimation using RGB-D camera," *Sensors*, vol. 19, no. 6, pp. 1287, 2019.

[18] L. Fletcher, G. Loy, N. Barnes, A. Zelinsky, "Correlating driver gaze with the road scene for driver assistance systems," *Robotics and Autonomous Systems*, vol. 52, no. 1, pp. 71-84, 2005.

 

[19] T. Wah, R. Gopal Raj, “Reservoir computing based echo state networks for ventricular heart beat classification,” *Applied Sciences*, vol. 9, no. 4, pp. 702, 2019.

[20] K. Kircher, C. Ahlström, “Issues related to the driver distraction detection algorithm AttenD,” *First International Conference on Driver Distraction and Inattention*, Gothenburg, Sweden, 2009.

[21] A. Tascillo, R. Miller, “An in-vehicle virtual driving assistant using neural networks,” *Proceedings of the International Joint Conference on Neural Networks*, Portland, OR, pp. 2418-2423, 2003.

[22] X. Fan, F. Wang, D. Song, Y. Lu, J. Liu, “GazMon: eye gazing enabled driving behavior monitoring and prediction,” *IEEE Transactions on Mobile Computing*, 2019.

[23] M. Hossin, M. Sulaiman, “A review on evaluation metrics for data classification evaluations,” *International Journal of Data Mining and Knowledge Management Process*, vol. 5, no. 2, pp. 1-11, 2015.