Methodology for in-the-Wild Driver Monitoring Dataset Formation

Alexandr Bulygin, Alexey Kashevnik
St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)
Saint Petersburg, Russia
alexandr_bulygin@mail.ru, alexey.kashevnik@iias.spb.su

Abstract—**Driver distraction and fatigue have become one of the leading causes of severe traffic accidents. Hence, the systems that implement driver monitoring systems are crucial. Usually such systems used a monocular camera to recognize driver behavior. Even with the growing development of advanced driver assistance systems and the introduction of third-level autonomous vehicles, this task is still trending and complex due to challenges such as in-cabin illumination change and the dynamic background. To reliably compare and validate driver inattention monitoring methods a limited number of public datasets are available. The paper proposes a methodology for in-the-wild dataset creation of vehicle driver for recording an oculomotor activity, a video images of a driver as well as relevant smartphone sensors that track vehicle movement. Based on the methodology we plan to conduct in-the-wild experiments.**

I. Introduction

Transportation plays a vital role in individual and social welfare, the economy, and quality of life. Its benefits, however, are not a free lunch. Society pays in terms of money (for vehicles' purchase, operational, and maintenance costs), social and ecological costs (resource utilization, exhaust and noise pollution, traffic jams), fatal or harmful traffic accidents, and so on [1].

Fatigued driving is one of the main causes of traffic accidents. Research shows that the probability of traffic accidents caused by fatigued driving is five times higher than normal driving. The annual traffic accidents caused by fatigue driving account for about 20% of the total accidents, accounting for more than 40% of the serious traffic accidents [2].

According to the global status report on road safety conducted by the World Health Organization (WHO) in 2015, 1.25 million traffic-related fatalities occur annually worldwide, with millions more sustaining serious injuries and living with long-term adverse health consequences; road traffic injuries are currently estimated to be the leading cause of death among young people, and the main cause of death among those aged 15–29 years. Road safety perception cannot be detached from the analysis of the driver behavior as the major part of traffic accidents is caused by human factors as it was inferred that they took part in the manifestation of 95% of all accidents [3].

If the current tendency lasts for a decade, an increased rate as high as 60–70% of road accidents could make it the 5th main cause of death by 2030. In monetary terms, the costs involved in road accident damages are estimated at more than half a trillion USD, which makes nearly 2% of the gross national product (GNP) of advanced economies, 1.5% of GNP of medium-income countries, and 1% of GNP of low-income countries [1].

Therefore, it has a great practical significance to monitor driver's fatigue status as well as response warning information in time.

Many studies have focused on modeling driver behavior, either for commercial purposes, management functions, or awareness campaigns. Their main goal is to explain the correlation between driver behavior and other factors through their model. It is a complex system characterized by a wide variety of variables and it has been proven that the majority of accidents are caused by human errors such as conscious law violations, distraction, inattention, fatigue, etc. The evolution of this area of study is made possible thanks to the progress of data analysis methods over the years. The development of these approaches improved the quality of driver behavior analysis and opened the door for new fields of applications [4].

The paper proposed a methodology for in-the-wild driver monitoring dataset formation that includes such tracking parameters as a driver's face and body image, data from eye tracker as well as vehicle telemetry. In the scope of the methodology we defined the sequence of actions that should be implemented for the dataset formation starting from the selection of participants for the experiment and the actions they perform to the sensors used and the analyzed parameters of the participants. The presented methodology has been tested in-the-wild to most accurately correspond to reality. In any case, the methodology created at the initial stage will be changed and supplemented during the experiment. Operation in-the-wild is complex and unpredictable, so it is impossible to take into account all possible factors and risks.

The structure of the paper is as follows. In Section II we present the related work in the topic of methodology for dataset preparation of vehicle drive. Section III describes the proposed models and algorithms of driver drowsiness detection. Section IV describes the proposed methodology that is based on the considered related work analysis. Section V describes the testing of the proposed methodology on a PC operator. The conclusion summarizes the paper.

II. RELATED WORK

In the paper [5], the method predicts the kinematics of vehicles that cover coverage patterns covering various driving patterns in normal driving. The online monitoring scheme is designed using an exponentially weighted moving average (EWMA) and a cumulative sum (CUSUM) chart that detects abnormal average lateral speeds and lane position prediction errors to warn of distracted driving.

In the paper [6], the dataset is mainly composed of two sets: the first one recorded in the daytime and the second one at nighttime. Each set consists of two synchronized data modalities, both from frontal and side views. More than 60 drivers are asked to execute 16 in-vehicle actions under a wide range of naturalistic driving settings. Dataset presents multiple modalities, spectrums, and views under different time and weather conditions.

In the paper [7], the study aims to optimize the Long Short-Term Memory (LSTM) model for phone usage detection based on vehicle dynamics sensor data from Shanghai Naturalistic Driving Study (SHNDS), China. A total of 1244 phone use events were extracted from videos of SH-NDS and analyzed against the focus driving baseline. Performance attributes included speed, longitudinal acceleration, lateral acceleration, lane offset, and steering wheel rate. Their mean, standard deviation, and predicted error (PE) were calculated and derived from 15 indicators. A Bidirectional layer and attention mechanism were added to the LSTM model for higher accuracy.

In the paper [8], the algorithm for the detection of drivers' manual distraction was proposed. The detection algorithm consists of two modules. The first module predicts the bounding boxes of the driver's right hand and right ear from RGB images. The second module takes the bounding boxes as input and predicts the type of distraction. 106,677 frames extracted from videos, which were collected from twenty participants in a driving simulator, were used for training (50%) and testing (50%).

The paper [9], provides an overview of driver distraction, then presents the available datasets and explores various signals for driver distraction analysis. After that, two forms of driver distraction (visual distraction and manual distraction) are analyzed separately.

In the paper [10], two generalized linear mixed models, one with at-fault safety-critical events (SCE) and the other with all-cause SCEs as the outcomes, was developed to compare the odds associated with common distraction types using data from the SHRP2 naturalistic driving study. Results: Adjusting for the environment and driver variation, 6 of 10 common distraction types significantly increased the risk of at-fault SCEs by 20-30%. The three most hazardous sources of distraction were handling in-cabin objects (OR = 14.3), mobile device use (OR = 2.4), and external distraction (OR = 1.8). Mobile device use and external distraction were also among the most commonly occurring distraction types (10.1% and 11.0%, respectively).

In the paper [11], naturalistic non-essential task involvement (NEST), a subset of SHRP2 data, was used to analyze non-essential task involvement and off-path (not in the direction of travel) gaze. In addition to assessing their relationship to environmental requirements, the age of the driver and the selected speed were taken into account. Results. Environments with higher visual complexity (characterized as visually complex and/or low visibility) were associated with a reduced likelihood of performing a secondary task, as well as a decrease in off-path looks, especially longer ones (>2 s). Drivers aged 35 and older were less likely to glance to the side than younger drivers. An increase in speed was associated with a decrease in the likelihood of completing the task under conditions of higher motor control complexity (characterized as poor surface conditions and/or a curved road), but not under conditions of lower complexity.

So, there is no dataset of oculomotor activity of drivers, together with video images of the driver as well as vehicle telemetry in-the-wild. Therefore, it is necessary to record such a dataset yourself. Driving in-the-wild depends on a large number of factors. Creating such a dataset in-the-wild is a very difficult task. Therefore, the creation of a methodology for creating a such dataset is a very urgent task.

III. FATIGUE DETECTION

We presented the reference model of driver drowsiness detection in detail in our paper [12]. We highlighted parameters that we are used for discussed in the paper methodology (see Fig. 1). Based on the static parameters, we obtain dynamic parameters, which can be calculated using computational parameters. Then we can detect fatigue, drowsiness, or loss of concentration. For example, consider temperature. If the temperature has dropped by 1 degree, then this indicates drowsiness.

The proposed algorithm for determining fatigue by eye movement (Fig. 2). If a person has extra-long gaze fixations, then this indicates fatigue. Also, extra-long gaze fixations may indicate the appearance of tunnel vision. The trajectory of gaze is associated with randomness and orderliness. If a person has a chaotic gaze, then this indicates fatigue. The moving gaze reflects the change in the tracing characteristics of the gaze, which suffer from fatigue. If the gaze moves in jumps, then this indicates fatigue. Two examples of eye movements indicative of fatigue are presented in Fig. 3. On the left half of the figure is an example of tunnel gaze, on the right half is an example of a chaotic gaze.

PERCLOSE is calculated automatically by Drive Safely system. The system takes images of the driver head for two seconds. Based on these images the system detects if the eye is opened or it is closed. After that, it calculates the number of closed eyes and divides it into the overall number of frames.

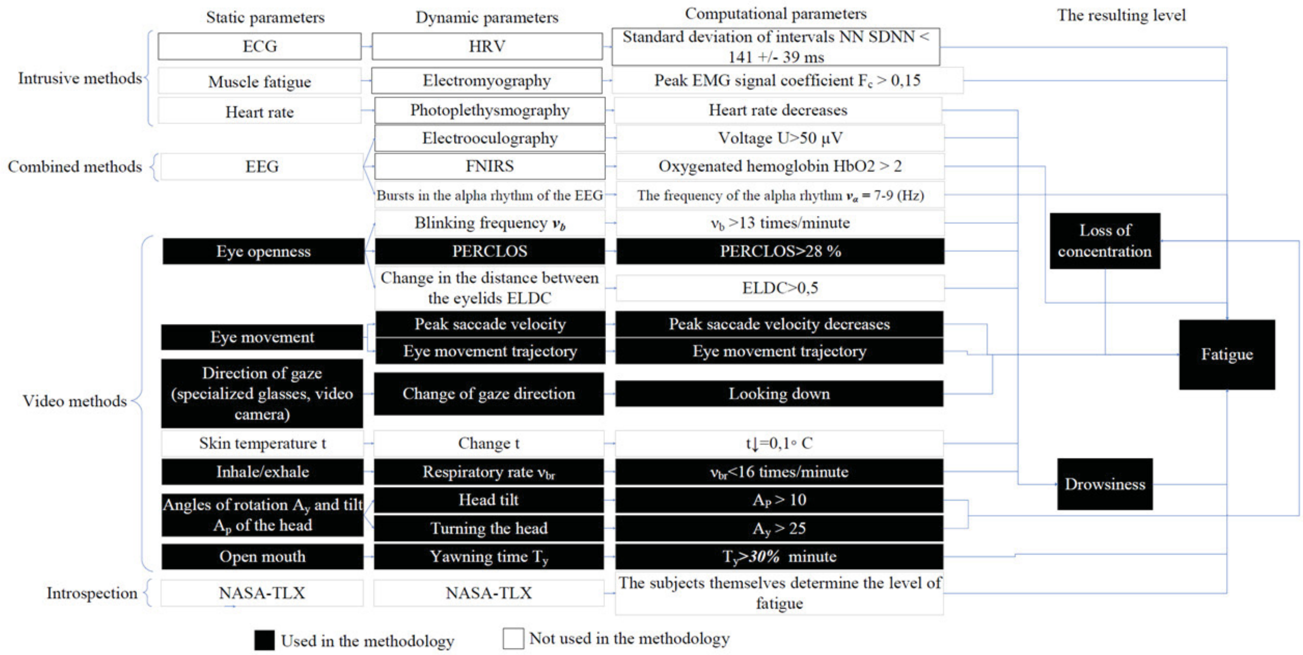

Fig. 1. Fatigue reference model

We proposed to calculate the distribution areas of gaze fixation points (circle) and saccades (lines). They are determined by the length of the trajectory of gaze movement. A circle with a radius of 100 pixels is drawn with the center at the gaze fixation point, in which the gaze moves. This radius value was chosen empirically. If the gaze is in this circle for more than 3 seconds, then an extra-long fixation is detected and fatigue is registered. If the eye goes beyond this circle, then a new circle and a line are drawn that connects the centers of the old and new circles. We propose to draw three circles. The circle in which the eye is currently located and 2 previous circles and lines between them. The trajectory of gaze movement is calculated by the formula (1).

$$traj = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, \quad (1)$$

where (x_1, y_1), (x_2, y_2) - coordinates of successive gaze fixation points. The speed is calculated according to the the formula (2).

$$V = \frac{S}{t}, \quad (2)$$

where V - eye movement speed, S - the length of the path traveled by the gaze in time t, t = 0.5 seconds. S is calculated using the formula (3).

$$S = S_c + S_s, \quad (3)$$

where S_c - length of the path traveled by the gaze in a circle, S_s - the length of the path traveled by the eye between the circles. Gaze speed is updated 1 time every 0.5 seconds.

I. Methodology

We developed the methodology for dataset formation of oculomotor activity of drivers together with video images of the driver and other relevant sensors to automate recording the data in-the-wild environment.

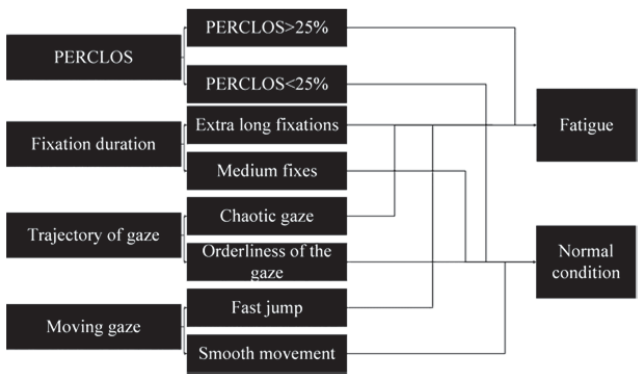

Fig. 2. Algorithm for determining fatigue based on eye movement

A. General Description

The developed methodology allows creating a dataset of a driver oculomotor activity (Fig. 5). We identified the following main phases: (1) preparation phase; (2) before driving phase; (3) during driving phase; and (4) after driving phase.

During the preparation phase participants first take a Gottschaldt test (see Section IV, B) to detect field dependence-field independence. We decided that the experiment requires 9 field-dependent and field-independent men and 3 field-dependent and field-independent women. Also, the list of maneuvers should be determined in this phase. We decided that the travel time should be approximately 50 minutes.

Fig. 3. Eye movements indicative of fatigue. Tunnel gaze (left), chaotic gaze (right)

During the driving phase the driver should estimate his/her degree of fatigue based on Karolinska Sleepiness Scale. While driving, the indicated maneuvers must be performed. Eye tracker from Pupil Labs measures the gaze direction as well as obtain images of the road and the direction of the driver's gaze. We used the developed earlier platform to collect the vehicle telemetry data. To record the data in the vehicle cabin we used Drive Safely mobile system [13] developed for Android-based smartphones. The system is a driver assistant and monitoring system which is responsible to detect dangerous situations in vehicle cabins and provide recommendations to the driver as well as collect all information got to the cloud server [14]. Also, Drive Safely allows to detect such parameters as PERCLOS, respiratory rate, angles of rotation and tilt of the head, mouth openness using the incabin camera. Using the camera installed in the car, such parameters of the driver as are detected.

During the after driving phase the driver also estimates his/her fatigue level based on Karolinska Sleepiness Scale. We show an example for dataset formation in the vehicle cabin in Fig. 6.

B. Gottschaldt test

The Gottschaldt test is carried out to detect field dependence and field independence (Fig. 7). This test detects individual differences in cognitive activity, the degree of freedom from external referents, or, in other words, the degree of orientation of a person when making decisions to his/her knowledge and experience, and not to external guidelines if they conflict with his experience.

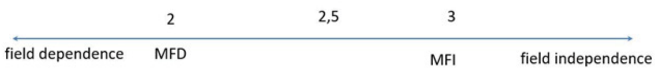

Fig.4. Field dependence-field independence scale

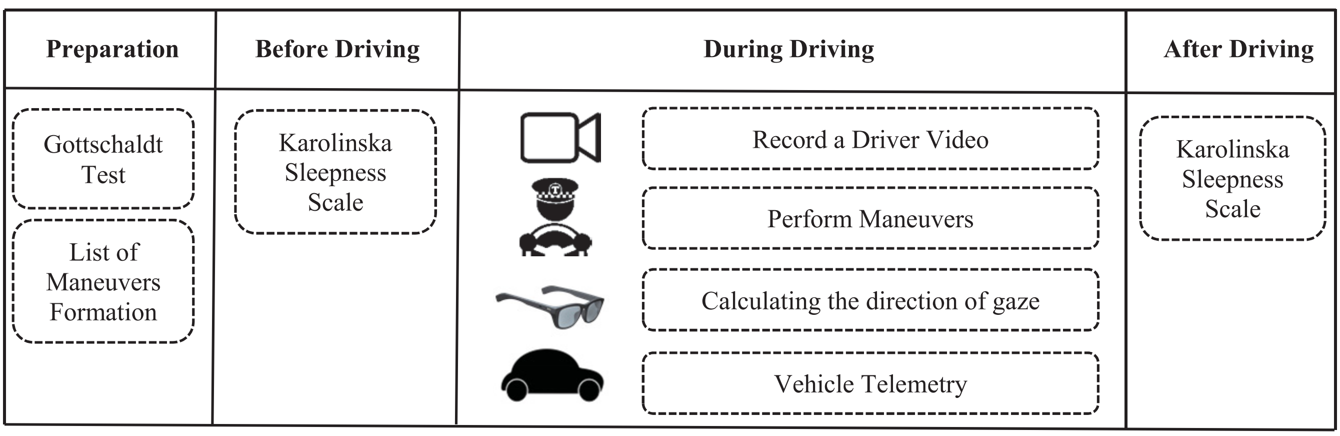

Fig. 5. Proposed Methodology

We highlighted the following characteristics of field dependence and field independence people. For field dependence people.

- More sociable, like social contacts.
- Choose a kind of occupation in which the means of activity are predetermined, stipulated, prefer the collective performance of the task.
- Are more prone to all sorts of illusions of perception.
- Synthetic perception.
- Eye movements.
    - Global image analysis (general to specific).
    - More scatter of fixation points on the image.

Characteristics of field independence people.

- More successful in intellectual activity.
- Choose a field of activity that requires high independence in the means of achieving the goal.
- The perceived "picture" is much more structured than field-dependent ones.
- The perceived object, the qualities of the object are perceived independently, separately from other objects, the qualities of these objects, are perceived simultaneously with this object.
- Analytical perception.
- Eye movements:
    - Local image analysis (from the particular to general).
    - Less scatter of gaze fixation points on the image.

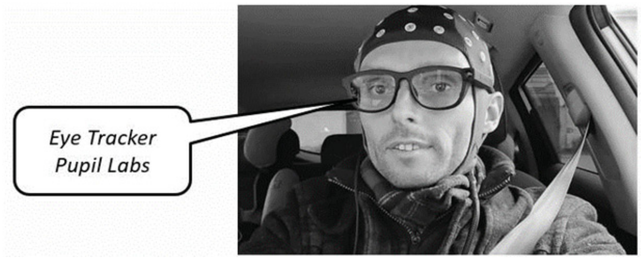

Fig. 6. Dataset formation in vehicle cabin

C. Karolinska Sleepness Scale

As mentioned before the experiment requires 9 field-dependent and 9 field-independent men and 3 field-dependent and 3 field-independent women. Before driving as well as after driving every driver should estimate his/her degree of fatigue using the Karolinska Sleepiness Scale to detect the degree of their fatigue. We show Karolinska Sleepiness Scale in Table I.

D. Maneuvers and route

While driving, the subject performs the following maneuvers:

- Crossroads (from 10 pieces)
- Pedestrian crossings (from 10 pieces)
- Spreads (from 3 pieces)
- Changeovers (on the freeway 5 pieces)
- Detour, overtaking (on the freeway 5 pieces)
- Parking (1 piece)
- City roads up to 60 km/h, freeways 110 km/h

Travel time is approximately 50 minutes. Below is an example of a route starting from the SPC RAS (Fig. 7).

II. EXPERIMENTS

Before testing on the driver, the methodology was tested on the PC operator since it is more simple task. Below in Fig. 8 is an example of calculating oculomotor parameters for a PC operator. Here the area of gaze fixation points, gaze movement trajectories, current gaze movement speed, the average speed for all videos, the minimum speed for all videos, maximum speed for all videos are calculated.

TABLE I. KAROLINSKA SLEEPINESS SCALE

Description	Point
Extremely alert	1
Very alert	2
Alert	3
Rather alert	4
Neither alert nor sleepy	5
Some signs of sleepiness	6
Sleepy, but no effort to keep awake	7
Sleepy, but some effort to keep awake	8
Very sleepy, great effort to keep awake, fighting sleep	9
Extremely sleepy, can’t keep awake	10

Fig. 7. Route from the Institute of St. Petersburg FRC RAS

Fig. 8. Calculation of oculomotor parameters for PC operator

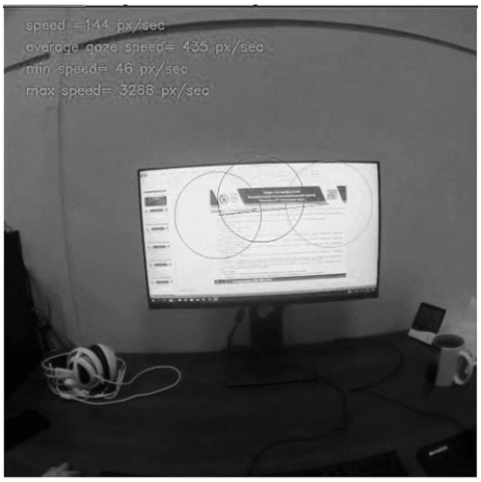

Fig. 9. Oculomotor parameters in a normal state

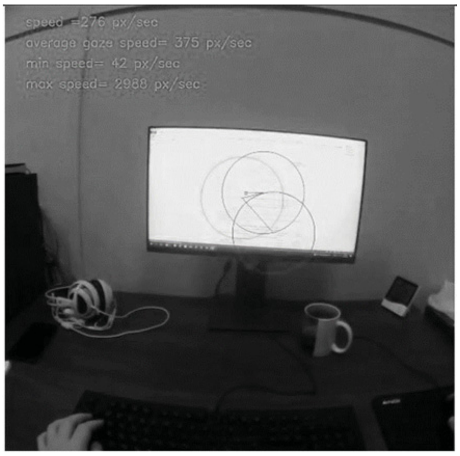

Fig. 10. Oculomotor parameters in a fatigue state

We show in Fig. 9 examples of calculating oculomotor parameters in the morning in a state of alert and in the evening in a state of fatigue in Fig. 10. As can be seen, the average speed of eye movements in the evening is lower than in the morning. Recordings were made in the morning and evening for 4 days. As can be seen from Table II, the average values of the average, minimum, and maximum speed during the state of fatigue are lower than in the state of alert.

TABLE II. AVERAGE VALUES OF THE PARAMETERS IN THE STATE OF ALERT AND IN THE STATE OF FATIGUE

	Driver	**Average speed (px/sec)**	**Minimum speed (px/sec)**	**Maximum speed (px/sec)**
State of fatigue	1	375	42	2988
	2	357	36	2954
	3	430	40	3362
	4	268	20	4016
	Average value	357	34,5	3330
State of alert	1	435	46	3268
	2	351	38	3582
	3	493	48	3472
	4	373	38	4462
	Average value	413	42,5	3696

VI. CONCLUSION

Using Pupil Labs glasses, the gaze movements of the PC operator were recorded at the beginning of the working day and at the end of the working day, when fatigued. The average values of the average, minimum and maximum speed during fatigue are lower than in the state of alert. Using the camera installed in the car and mobile app Drive Safely, such parameters of the driver as PERCLOS, respiratory rate, angles of rotation and tilt of the head, mouth openness are detected. Using glasses for calculating the direction of gaze Pupil Labs obtain images of the road and the direction of the driver's gaze. Using GPS/GLONASS, speed and acceleration are determined.

For the future work we plan to record 20 videos in the state of alert and 20 videos in the state of fatigue. At the first stage, compared according to the following parameters: (1) average speed for all videos; (2) the minimum speed for all videos; (3) maximum speed for all videos. At the second stage we plan to compare according to the following parameters: (1) gaze fixation duration; (2) trajectory of gaze movement; (2) moving gaze.

ACKNOWLEDGMENT

Fatigue detection as well as experiments (Section III and Section IV) have been developed due to the work supported by the Russian Foundation for Basic Research, project number # 19-29-06099. The methodology (Section III) is due to Russian State Research, project number FFZF-2022-0005.

REFERENCES

[1] Khan M., Lee S. A Comprehensive Survey of Driving Monitoring and Assistance Systems, *2020 IEEE Region 10 Symposium (TENSYMP)*, Bangladesh, 2020, pp. 125-159.

[2] Liu Z., Peng Y., Hu W. Driver fatigue detection based on deeply-learned facial expression representation, Journal of Visual Communication and Image Representation, vol. 71, no.1., 2020, pp. 452-461.

[3] *Elamrani Z., Mousannif H., Moatassime H.* The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review // Engineering Applications of Artificial Intelligence. – 2020. – Vol. 87, N 5. - P. 336-352.

[4] Zinebi K., Souissi N., Tikito K. Driver Behavior Analysis Methods: Applications oriented study // The 3rd International Conference on Big Data, Cloud and Applications BDCA'18. – Morocco, 2018. – P. 224-240.

[5] Sun W., Aguirre M., Jin J. Online distraction detection for a naturalistic driving dataset using kinematic motion models and a multiple model algorithm, *Transportation Research Part C: Emerging Technologies*, vol. 130, no.1, 2021, pp. 213-233.

[6] Jegham I., Khalifa A., Alouani I. A novel public dataset for multimodal multiview and multispectral driver distraction analysis: 3MDAD, *Signal Processing: Image Communication*, vol. 88, 2020, pp. 166-174.

[7] Wang X., Xu R., Zhang S. Driver distraction detection based on vehicle dynamics using naturalistic driving data, *Transportation Research Part C*, vol. 136, no. 1., 2022, pp. 15-27.

[8] Li L., Zhong B., Hutmacher C., Detection of driver manual distraction via image-based hand and ear recognition, *Accident Analysis and Prevention*, 2020, pp. 110-121.

[9] Li W., Huang J., Xie G., A survey on vision-based driver distraction analysis, *Journal of Systems Architecture*, vol.121, 2021, pp. 78-94.

[10] Liang O., Yang C., Determining the risk of driver-at-fault events associated with common distraction types using naturalistic driving data, *Journal of Safety Research*, 2021, pp. 45-50.

[11] Risteska M., Kanaan D., Donmez B. The effect of driving demands on distraction engagement and glance behaviors: Results from naturalistic data, *Safety Science*, vol.136, 2021, pp. 946-960.

[12] Bulygin A., Kashevnik A., "Image-Based Fatigue Detection of Vehicle Driver: State-of-the-Art and Reference Model," 2021 30th Conference of Open Innovations Association FRUCT, 2021, pp. 24-31.

[13] Kashevnik A., Lashkov I., Gurtov A. Methodology and Mobile Application for Driver Behavior Analysis and Accident Prevention. IEEE Transactions on Intelligent Transportation Systems, IEEE. 2019. Vol. 21(6). P. 2427–2436.

[14] Kashevnik A., Lashkov I., Ponomarev A., Teslya N., Gurtov A. Cloud-Based Driver Monitoring System Using a Smartphone. IEEE Sensors, IEEE. 2020. Vol. 20(12). P. 6701–6715.