International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 2, April 2024, pp. 2212~2225
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i2.pp2212-2225

Analysis of driving style using self-organizing maps to analyze driver behavior

Yulia Shichkina, Roza Fatkieva, Maxim Kopylov
Department of Computer Science and Engineering, Saint Petersburg Electrotechnical University LETI, Saint Petersburg, Russia

Article Info

Article history:

Received Feb 12, 2023
Revised Oct 12, 2023
Accepted Dec 13, 2023

Keywords:

Dangerous driving
Driving style
Kohonen self-organizing maps
Obd II
Vehicle control

ABSTRACT

Modern life is strongly associated with the use of cars, but the increase in acceleration speeds and their maneuverability leads to a dangerous driving style for some drivers. In these conditions, the development of a method that allows you to track the behavior of the driver is relevant. The article provides an overview of existing methods and models for assessing the functioning of motor vehicles and driver behavior. Based on this, a combined algorithm for recognizing driving style is proposed. To do this, a set of input data was formed, including 20 descriptive features: About the environment, the driver's behavior and the characteristics of the functioning of the car, collected using OBD II. The generated data set is sent to the Kohonen network, where clustering is performed according to driving style and degree of danger. Getting the driving characteristics into a particular cluster allows you to switch to the private indicators of an individual driver and considering individual driving characteristics. The application of the method allows you to identify potentially dangerous driving styles that can prevent accidents.

Corresponding Author:

Yulia Shichkina
Department of Computer Science and Engineering, Saint Petersburg Electrotechnical University LETI
ul. Professora Popova 5, 197022 Saint Petersburg, Russian
Email: strange.y@mail.ru

1. INTRODUCTION

The accelerated development of road transport creates serious problems for society in terms of social, economic and environmental costs. The growing trend towards an increase in the level of urbanization leads to an increase in the density of car traffic in cities, while the quality of driver training does not always correspond to the proper level. All this leads to an increase in accidents involving road transport. According to the World Health Organization (WHO), road traffic crashes cause 1.3 million annual deaths and 50 million injuries worldwide [1]. At the same time, the relationship between driver behavior and driving style can be one way to predict and reduce accidents. This relationship should take into account the current traffic situation.

The increase in acceleration and maneuverability of modern vehicles leads to the fact that drivers develop a dangerous driving style, often not conscious of them. Under these conditions, the use of a traffic analysis system allows to monitor the behavior of the driver and warn of his aggressive behavior. Taking into account the specifics of driver behavior and current traffic, the system can create recommendations for the driver. These actions of the system will help to make traffic on the roads safer and improve a person's driving skills.

Over the past decades, various methods and models have been proposed to describe and evaluate the traffic situation. In study [2], three basic elements of the driving process were studied: The driver, the vehicle, the road environment and their interaction. It is shown that the system of monitoring and driving assistance improves the quality of driving by collecting and processing data received from sensors.

In study [3], [4], a framework for driving style analysis based on the Bayesian non-parametric approach is presented. Wang *et al.* [3] using a hidden semi-Markov model to extract driving patterns that uses multivariate time series traffic data, providing a semantically interpretable way to analyze driver behavior and driving style. Han *et al.* [4] presented method uses the Euclidean distance of each pair of elements in the feature vector, classifying the driving style into seven levels from normal to aggressive.

The problem of vehicle collision recognition is considered in [5]. The method uses the characteristics of the vehicle trajectory, including acceleration, relative velocity, and relative distance. Discrete Fourier transform (DFT) and discrete wavelet transform (DWT) are used for processing. The use of the support vector machine over the vector of input characteristics showed that the method outperforms others with an accuracy of 91.7% due to the extraction of objects by the method of DWT.

The study [6] shows the relationship between driving style and scenario and the fuel consumption of a conventional vehicle. It is shown that fuel consumption for conventional vehicles is 13% lower in quiet city driving, but almost 34% higher in unstable highway traffic compared to standard driving cycles. If the driving cycle is predicted and used to adjust the power distribution in a hybrid electric vehicle, then its fuel consumption can be reduced by up to 12% in urban driving.

There are works devoted to the consideration of the driving style of unmanned vehicles. Articles [7], [8] consider various strategies for the behavior of autonomous vehicles in a changing environment. Studies [7], [9] are devoted to the assessment of trust in the exchange of information between vehicles. Studies [10]–[12] continue the topic of assessing driving style, but in the key of its impact on traffic safety. The indicators characterizing the driving style in [10], [11] are ranked according to a hierarchical structure and allow to identify ambiguity in the driver's thinking style, which affects safety. The multilevel models of driving style assessment described in [12] take into account behavioral aspects (driver's emotions and driver's condition) to ensure greater safety and comfort. Machine learning methods for driving style classification are reflected in the works [13]–[20]. The work [13] shows the expediency of using the support vector machine to identify the driving style. In [15], a driving style classification method based on constructed working driving charts with driving style assessment using a convolutional neural network (CNN) is presented. The use of a recurrent neural network for processing data obtained from global positioning system (GPS) is presented in [17]. In study [18], driving style is assessed using principal component analysis, using a fuzzy c-means clustering algorithm and a parameter identification algorithm based on the support vector machine.

Works [21]–[23] are devoted to clustering problems. In study [21], driving style identification is carried out using self-organizing Kohonen maps using the frequency distributions of the identified maneuvers, which are then interpreted to improve fuel consumption. Zepeda *et al.* [22] proposed to use the apparatus of dynamic clustering, in which clusters change depending on the changing environment. In [23], a method of collective learning with a majority of votes is proposed.

The application of the game-theoretic approach is proposed in [24], [25]. In study [24], an approach based on the Bayesian Nash equilibrium (BNS) game strategy for recommending optimal vehicle maneuvers in real time is considered. It was shown in [25] that, compared with the Nash equilibrium approach with a normal driving style, the cost of decision-making using the game-theoretic Stackelberg approach is reduced by more than 20%.

The methods of classifying driving style presented in Table 1 have shown that the assessment of drivers' behavior when driving vehicles is of great importance for the road safety system. However, despite numerous studies, developers still face various problems, including insufficient model accuracy, a large number of learning parameters and instability of classification results, as shown in Table 1. There are also problems in the assembly and processing methods of the received data, since they often turn out to be incomplete. As can be seen from Table 1 of the presented studies, the classification by driving style is most often based on an analysis of the functioning of a motor vehicle (8 studies), then, according to the degree of influence, methods for assessing the psychophysiological state of the driver follow (6 studies) and only 2 studies are aimed at assessing the totality of all three factors, including an assessment of the road situation.

Thus, the main problem is to assess the degree of influence of the dynamics of environmental change on the dynamics of driving style change. Identification of the most significant parameters of the environment and driving style will allow to form individual recommendations to the driver on the road. To solve this problem, we have developed a combined method for the selection of vehicles in the flow with the possibility of grouping their main characteristics of functioning into clusters. These clusters reflect driving style and identify the main features of dangerous driving. The developed method is presented in this paper.

The distinctive feature of the proposed method is the application of image processing algorithms together with the Kohonen network. With the help of image processing algorithms, we identify the "visual field of the vehicle ahead". With the help of the Kohonen network we cluster the main characteristics reflecting the driving style based on the obtained images and car performance characteristics.

The novelty of the approach presented in the paper is to consider the relationship between driver behavior and the environment by analyzing: i) a set of characteristics of the environment in the form of identification of objects, with the assessment of changes in their location by coordinates in space and ii) a set of characteristics of vehicle functioning. Deviation from the threshold values of both sets characterizes a dangerous driving style. The proposed approach demonstrates a deeper look at the dynamics of driving behavior under the influence of the environment.

Table 1. Comparative analysis of existing methods

Method	Assessment of the mental and physical condition of the driver	Assessment of the dynamics of changes in the situation	Vehicle condition monitoring	Classification by driving style
The Bayesian nonparametric learning method based on a hidden semi-Markov model (HSMM) [3], [4]	+	+[3]/-[4]	+	+
Vehicle collision recognition using DFT, DWT and statistical method [5]	-	-	+	+
Classification of driving based on the driver's thinking style [10]–[12]	+	-	+[12]/-[10], [11]	+
Classification method based on support vectors [13], [18], [19]	+	+[17]/-[18], [19]	+	+
Classification method based on neural network multi-layer perceptron (MLP) [14], CNN, long short-term memory (LSTM) network, and pre train-LSTM [15], [17], [20], random forest [16]	-[14], [17], [20]/+[15], [16]	-	+	+
Application of self-organizing maps [21]			+	+
Classification method, which is enabled by ensemble learning [23]	-	-	+	+
Gaming model [24], [25]	+		+	+

2. METHOD FOR RECOGNIZING DRIVING STYLE BASED ON A SELF-ORGANIZING KOHONEN MAP

The ability to recognize the driving style depends on the characteristics of the vehicle, the parameters and data transmission channels between its elements, the psychophysical characteristics of the driver, as well as the dynamic change in the current state of the environment. It is assumed that the strategy of behavior when driving a motor vehicle is implemented in a way that ensures the maximum achievement of the goal pursued by the driver. In order to take timely measures to ensure road safety, it is necessary to determine possible scenarios for the behavior of drivers in a dynamically changing environment.

The purpose of the study is to identify features that can classify the driving style of drivers, taking into account various aspects of the traffic situation. The study aims to identify key characteristics of driver behavior that can be used to categorize their driving style depending on current road conditions. This approach provides a more accurate and adaptive classification of drivers, improving understanding of the impact of the road environment on their driving behavior. However, it is not always possible to apply the currently existing methods of assessing the road situation. For example, the method of estimating the parameters of the traffic flow in real time from video from unmanned aerial vehicles is difficult to apply to correct driving style due to delays introduced during signal transmission [26]. The method of assessing the road situation [27] is based on training on a set of stereo image data and does not allow dynamically making changes to the training data sets. The methods [28], [29] are based on lidar readings and cannot always objectively take into account the temporal stability of variable states.

To effectively solve the problem, taking into account time delays, it is recommended to use clustering methods, such as Kohonen networks [30]. This method allows you to group similar input data by quantitative characteristics. The use of the clustering device provides a more accurate and systematic selection of similar objects, which is useful when analyzing dynamic data in driving style recognition tasks. In this case, we can consider a set of objects-vehicles as [31]:

$$S = \{s_1, s_2, \dots, s_l\},$$

where s_i is analyzed vehicle. The behavior of each driver of which is characterized by the vector.

$X = \{x_1, x_2, \dots, x_n\}.$

There is a need to build a set of clusters C and map a F set S to a set C, that is, $F: S \rightarrow C$. The task of clustering by type of driving is to build a set:

$$\mathrm{C} = \{\mathrm{c}_1, \mathrm{c}_2, \dots, \mathrm{c}_k, \dots \mathrm{c}_M\},$$

where c_k, is a cluster containing similar driving style objects from the set S:

$$c_k = \{s_i \in S \; and \; d(s_i, s_j) < h\},$$

h is a value that determines the measure of proximity for the inclusion of objects in one cluster, d_{ij} is a measure of proximity between objects, called distance.

The measure of proximity is a key characteristic. Because if the distance is d_{ij} less than a certain value h, then the driving styles are considered close and are placed in one cluster. Otherwise, the driving styles are considered to be different from each other and they are placed in different clusters.

Dynamic changes in driving behavior, including braking, moving and stopping a vehicle on the road, require a corresponding change in the driving style of the vehicles behind. This leads to the classification of driving style taking into account its influence on the behavior of all road users. However, predicting the behavior of drivers only by changes in the road environment does not reflect the internal processes occurring in the functioning of the vehicle (e.g., sharp acceleration at short distance intervals). Under these conditions, it is reasonable to use an algorithm for recognizing driving style based on the assessment of environmental changes and vehicle functioning characteristics collected, for example, using OBD-2.

To solve the problem of assessing the environment and responding to it, it is possible to use stereo sensors or cameras that continuously collect images of the changing environment while driving on the road, allowing the generation of data vectors that uniquely identify changes in the road situation. This recognition process includes semantic segmentation, map generation and object detection. For this purpose, we propose an approach that consists of the following steps as shown in Figure 1.

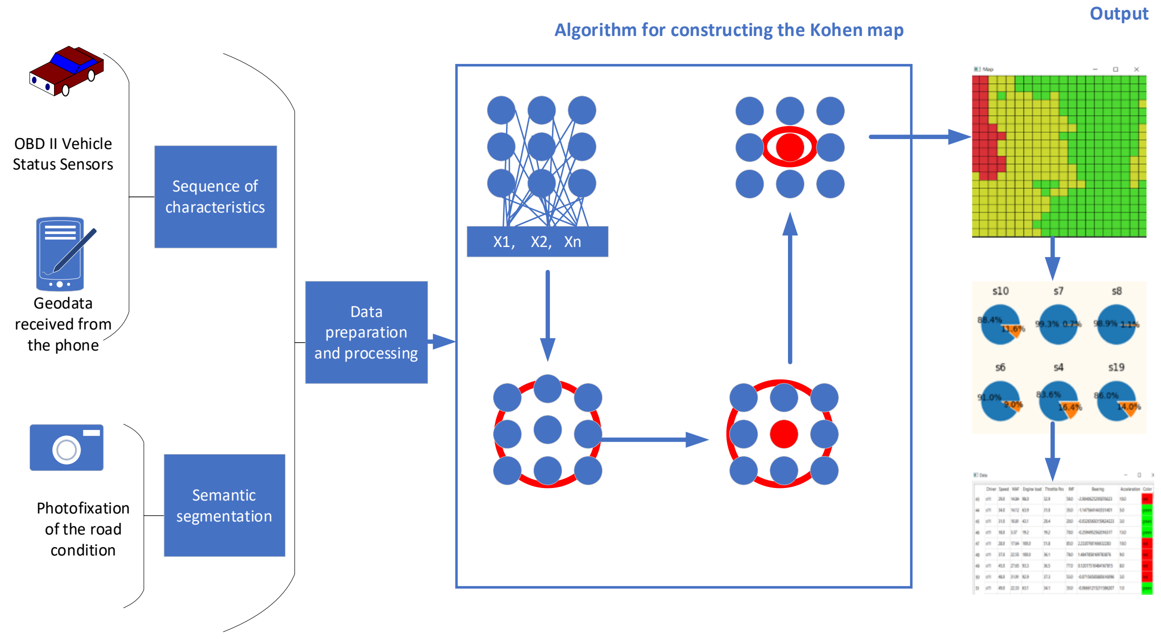


Figure 1. A combined method for recognizing driving style

a. A set of images of the current traffic situation is formed to identify moving vehicles.
b. The resulting image is divided into N pixels, which allows further identification of nearby pixels belonging to the object.
c. A set of objects is formed in the resulting image using the Detectron2 library, which provides the most advanced algorithms for detecting and segmenting objects in a photo.

d. A set of characteristics is formed that identifies the vehicle and determines the road situation based on the obtained images. An example is highlighting all available contours of vehicles on the map with the identification of their area in pixels, with an estimate of the ratio of the image area to the area of the contour of vehicles. To calculate the contour of a vehicle, the ratio $P = M/N$ is used, where N is the total number of pixels in the photo; M is the sum of all unique pixels inside the boxes that characterize each car in the photo. Obviously, P belongs to the interval [0, 1]. This parameter is normalized and does not require any processing. It characterizes the level that reflects the traffic situation and the traffic density of vehicles. If there are many, then the coefficient will be high, otherwise low.
e. Formation of a data set that includes vehicle density and vehicle performance information collected using OBD-2.
f. Construction of basic driving style characteristics. For this purpose, it is proposed to use a self-organizing Kohonen map on the obtained dataset. This approach allows to realize the task of visualization and clustering of the main characteristics reflecting the driving style, as well as to identify their feature sets.

The neural network contains 100 functional radial basis function (RBF) modules organized as a 20×20 square lattice, with the Euclidean function allowing to estimate the proximity of the distance between the estimated parameters, random initialization and the Gaussian function h. Note that the network operation algorithm includes five main processes shown in Figure 2. i) Formation of a set of characteristics and organization of data collection, including information about objects on the road and characteristics that reflect the functioning of the vehicle; ii) Preparation and processing of data; iii) Initialization of initial scales; iv) Learning; v) Evaluation of errors; and vi) Evaluation of the result.

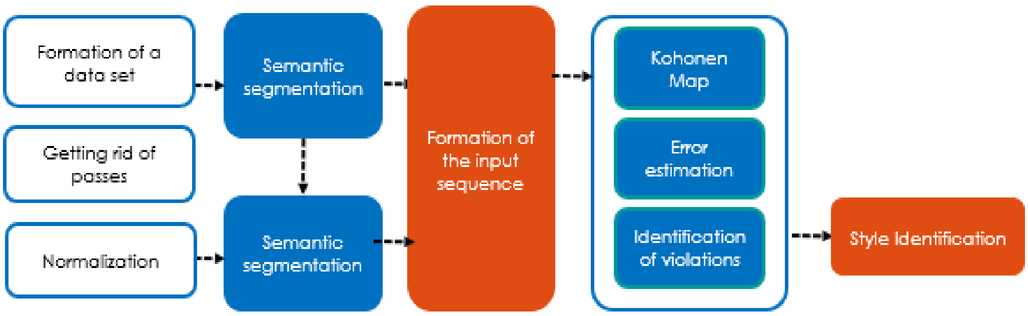


Figure 2. The method for recognizing driving style based on a self-organizing Kohonen map

Consider in more detail each of the stages of the formation of a self-organizing Kohonen map:

Step 1: At the stage of forming a set of characteristics and organizing data collection. it is advisable to select and form vectors from several indicators of the functioning of the vehicle (up to 50 features). However, with this approach, problems arise in which it is difficult to combine a significant number of indicators into groups that make it possible to identify interrelations and patterns between them that reflect the driving style. The use of transformations on the collected data using Kohonen maps makes it possible not only to reduce the dimension of the vector of indicators, but also to suggest relationships between indicators located in the same area and/or in closely spaced cells. This allows to define areas that are responsible for a particular driving style, for example, "dangerous driving". To do this, from the space of characteristics that reflect the functioning of the vehicle, those that can identify a driving style violation are selected. It is advisable to collect data using the standard for presenting digital data on the state of OBD II vehicle components and systems (for example, speeding, sudden braking or changing lanes). Based on these data, a data vector is formed, which is fed to the processing stage.

Step 2: Data processing consists in finding and removing outliers in the analyzed sequence. Additionally, possible errors are searched and corrected, followed by preprocessing and normalization of the data. This process aims to ensure the accuracy and reliability of the data before further analysis of driving styles.

The first step is to evaluate the errors obtained during data collection. The most common errors are missing data (null sequence) and getting unlikely data values (for example, the angle of rotation of the car by 180 on a small-time scale). Their finding, with the subsequent correction of the data, is carried out at the stage of data preprocessing. Data preprocessing, if necessary, allows to establish relationships between neighboring values and correct them in case of distortion. On the other hand, data preprocessing can be used to form a convolution of various indicators obtained during data collection, in this case, it becomes necessary to normalize data, which may be different in terms of the characteristics taken:

$$x_{iN} = \frac{x_i}{\sqrt{\sum_{i=0}^{m} x_i^2}},$$

where x_{iN} is the normalized component of the input vector; and $x_i(t)$ is value of the input vector of characteristics.

Step 3: Initialization of initial weights. This stage is quite important because it can significantly affect the learning process. For example, if the weights are set incorrectly, it is possible to be in the local minimum of the learning algorithm (i.e., the neural network will learn incorrectly for some cases with fixing this result). The optimal choice of weights is the choice of random weights from a given interval. In our case, we use randomly generated weights from 0 to 1, that allows to train the model most accurately with unknown input data.

Step 4: Learning process. Learning takes place in steps at each of the iterations t. During the training, the model adapts to new data and improves its ability to recognize driving styles. This stage is key to improving the efficiency and accuracy of the system, which ensures more accurate classification of drivers and higher resistance to changes in the road environment.

Step 4.1: The input vector is selected $x_i(t)$ from the complete set of input data X. To determine the "distance" between the input vector and the neuron $M(t)_c$, the usual Euclidean distance is used:

$$d(x_i, w_j) = \sqrt{\sum_{j=0}^{n} (x_i - w_j)^2},$$

where $w_j(t)$ weight is the node vector $M(t)_c$. The neuron with the minimum distance is assigned the nearest one:

$$d(x_i, w_j) = \underbrace{min}_{1 \le i \le n}(x_i, w_j)$$

Step 4.2: The neighborhood of the nearest neuron is calculated. To do this, we find all the neurons that lie no farther from our neuron than this distance. As a neighborhood function, it is possible to use the Gaussian function [32]:

$$h_{cj}(t) = \alpha(t) exp \frac{\|r_c - r_j\|^2}{2\sigma^2(t)},$$

where α is a training factor, monotonically decreasing with each subsequent iteration; r_c, r_j coordinates of nodes $M(t)_c$ and $M(t)_j$ on the map; σ a factor that reduces the number of neighbors with iterations. All neurons that fall into the specified area are subjected to a weight adjustment operation.

Step 4.3: Weight adjustment. The operation is necessary so that the weights of all nodes that are neighbors approximate the observation in question. For each neuron of the previous step, the operation is applied:

$$m_j = m_j(t-1) + h_{cj} \cdot x_i(t) - m_j(t-1).$$

Step 4.4: Refinement of training is a necessary step, since at the initial stage a sufficiently large value of the neighborhood of the nearest neuron is chosen, which allows one to arrange the vectors of neurons in accordance with the distribution of examples in the sample. Then there is a need to reduce this distance and the learning factor, proportionally with each iteration. To do this, it is advisable to use any decreasing functions, including the Gaussian function. To reduce the learning coefficient, it is possible to use a monotonic decrease of the form:

$$\beta(t) = \frac{A}{t},$$

where is the A initial value of the learning coefficient.

It is possible to check the completion of the training cycle by generating a small value of the error functional defined in step 6. This allows you to evaluate the effectiveness and accuracy of the trained model at the refinement stage. It also helps to fine-tune the model to the current data, providing high accuracy and stability in a dynamically changing environment.

Step 4.5: As soon as the iteration of the loop is completed, the transition to the next iteration occurs in step 4.1. To doc this, a new random input vector is selected and similar actions are performed for it. Learning occurs as long as the quantization error at the input vectors is large enough to estimate.

Step 5. Estimation of the map error, for which it is possible to use the arithmetic mean of the distances between the characteristics of the input vector and the weight vectors:

$$Q = \frac{1}{\mathrm{T}} \sum_{j=0}^{n} \quad (x_i - w_j)^2. \tag{1}$$

Step 6: Formation of information about driving style. The transition from vectors with a large number of indicators to a two-dimensional Kohonen map makes it possible to highlight zones, broken down by driving style and construct areas that can be characterized as dangerous and safe driving. This allows to check the vector for belonging to a dangerous area and identify the driving style at a given point in time.

Based on this approach, there are two ways to work with data and a neural network: i) Definition of dangerous driving in the moment. During the operation of the vehicle, with a given time interval, information is collected about the current state of the vehicle and the traffic situation, followed by checking the analyzed data for danger. In the event of dangerous driving, a driver warning is triggered: and ii) Determination of driving style. In this case, all data on the functioning of the vehicle are analyzed, all data with the identification of personal information. The resulting selection of vectors allows to identify which vectors are dangerous. Next, it is possible to consider the ratio of "dangerous" places in the studied vector to the total number of vectors, which makes it possible to set a threshold value that identifies dangerous driving and evaluate the driving style. If the value of the ratio is too high, then the driving style can be classified as dangerous.

3. PREPARING THE DATA AND CONDUCTING THE EXPERIMENT

In any observation, the preparation of data and the conduct of an experiment, carried out at all stages of the verification of the study, play an important role. The first stage usually includes preliminary processing of input data, during which data analysis and selection of informative features are carried out. The second stage is related to the modeling of the process under study. The third stage is to obtain statistical data and evaluate them. Based on this, the whole process of solving the problem can be divided into three big steps: i) pre-processing of the input data, ii) formation of the Kohonen map, and iii) obtaining group and individual statistical data on driving style.

For the study, data were taken that were publicly available from the Kaggle website [33], that allows extracting about 20 descriptive features of car functioning from the described dataset. In the generated training and test data, only those features were selected that can identify a violation of driving safety. Uninformative features are discarded, which are a constant value, as well as features that contain a large number of gaps. For recognition, the following performance indicators of the car were taken: i) the longitude-latitude pair in which the car is currently located; ii) speed; iii) engine load; iv) intake manifold pressure; v) mass airflow; vi) throttle position; and vii) acceleration.

During preprocessing, images with marked objects are fed to the software input. The JSON type file for each image contains a list of coordinates and classes of detected objects. This allows to form a vector of "car field of view space identification" for detecting moving objects.

Preparation of data for transmission to the neural network consisted in the formation of a vector consisting of the presented indicators taken in the time interval between measurements of 1 second. From the received dataset, gaps are removed and data is normalized. The implementation of the model and the learning process were carried out using the MiniSom library [34]. To train the model for further interpretation of the data, a Python program was developed that uses three classes: i) detector class for image processing; ii) loader class for processing, preparing input data of the data set into the neural model; and iii) minicom class for training the model based on the received input data and further processing and interpretation of information.

Figure 3 shows an example of practical application of the detector class. Its use allows to select images of objects moving in the flow (in the form of a frame of red color). For this purpose, all available contours of vehicles on the map are selected, their area is identified in pixels and the ratio of the image area to the area of the vehicle contour is estimated. This allows to generate a list of coordinates and classes of detected objects. The obtained data in the form of a JSON file generated for each image serve as a set of input data for the class loader, which then passes the data to the Kohonen network. Four groups of output data are then generated from the data for further analysis:

- Kohonen maps in Figure 4. The color indicates the areas in which the standard indicators of the vehicle movement process exceed the threshold value of the danger level: green is low risk; yellow is moderate risk; red is high risk of creating a dangerous situation.
- Kohonen maps with driving hazard degree. In these maps, the hazard degree is marked with a blue circle on the cluster in Figure 5.
- Pie charts for monitoring groups of drivers and individual participants Figure 6.
- Tables with obtained statistical data on the performance characteristics of a group of vehicles as shown in Figure 7.

The completed stages of data processing include the analysis and filtering of information, which makes it possible to visualize critical areas and move on to a more detailed consideration of the individual characteristics of the driving style. The data obtained after processing not only provides an overall picture of the dangerous style, but also allows you to highlight unique aspects of driver behavior. This approach to data analysis opens up new opportunities to identify individual driving characteristics and identify the most critical aspects of safety.

Figure 3. Example of vehicle identification for the transmission of images about the current traffic situation

4. RESULTS AND DISCUSSION

Driving style recognition and reckless driving classification in traffic conditions is a mapping of the similarity of different types of behavior under given conditions onto a map. If you arrange the location of the vector of indicators in certain cells of the map, then this will allow to form areas of dangerous driving. Figure 4 shows a map of the trained network. The color indicates the areas in which the standard indicators of the vehicle movement process exceed the threshold value of the danger level: green is low risk; yellow is moderate risk; red is high risk of creating a dangerous situation.

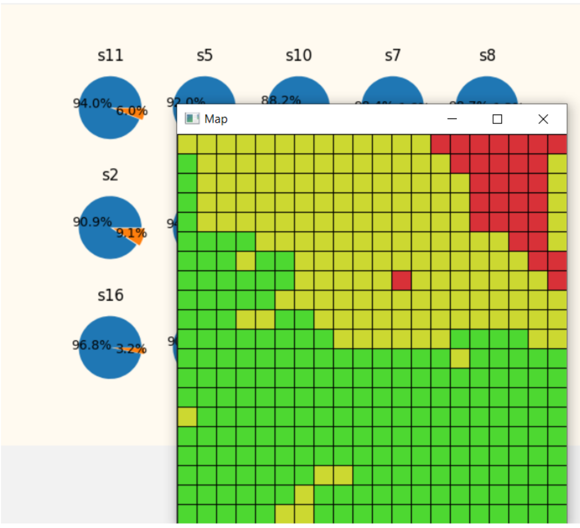


Figure 4. Example of the resulting Kohonen map

In the center of the figure is the Kohonen map obtained from training in Figure 5, showing the overall behavior of all vehicles in a given time interval. The hit of the driving characteristic in one or the cluster is canceled by a blue circle, while the thicker the circle, the more hits in a particular cell occurred. The absence of a circle means that there were no hits in this cell. This allows to identify the style of behavior for each individual driver. As can be seen, the s4 driver (on the right) exhibits more dangerous behavior compared to the driving style of the s7 driver on the left.

Thus, the presented Kohonen 's self-organizing map can serve as a basis for the further development of new, more complex models that recognize types of dangerous driving as a set of driver behavioral characteristics, which consist in organizing a danger for the movement of other vehicles. The subsequent development of the method can be aimed at expanding the set of indicators that identify dangerous driving, as well as automating the process of recognizing dangerous driving for individual groups of drivers. The application of the method for individual sections of the road, taking into account the time intervals of movement, makes it possible to identify sections that are dangerous for traffic, which makes it possible to correct the traffic situation on more dangerous sections of the road.

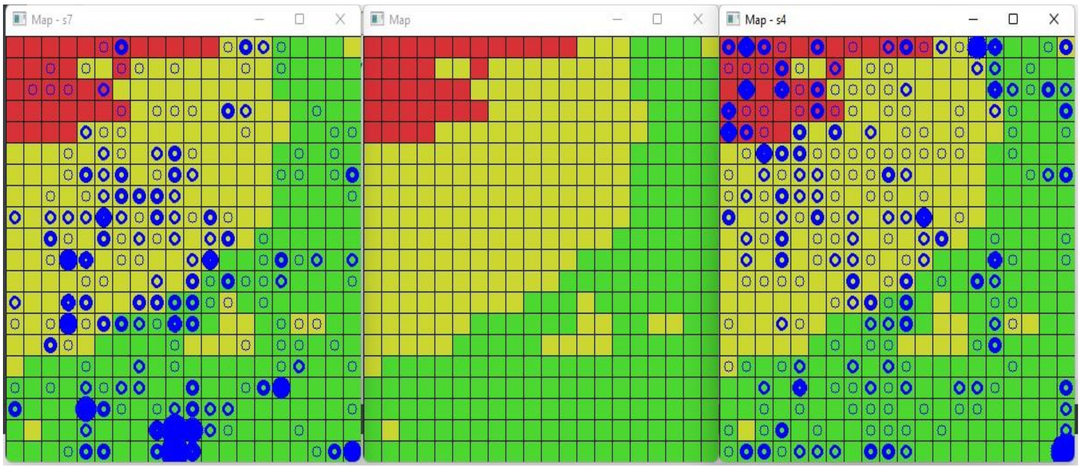


Figure 5. An example of the resulting map

The transition from the self-organizing map of Kohonen's driving style to the private indicators of an individual driver makes it possible to evaluate both group and individual driving characteristics. This makes it possible to rank drivers from the presented dataset from the safest driving style to the most dangerous. Visualization of data according to the degree of danger, for example, by indicating the percentage of lines that fell into dangerous cells and reflecting them using pie charts, allows to automate monitoring of both individual drivers in Figure 6(a) and the group of analyzed drivers in Figure 6(b). For the dynamic formation of a driving profile, it is necessary to provide for obtaining data for a long time period on a given space of indicators, this makes it possible to obtain an average driving style on the required section of the road, and to identify deviations in the driver's behavior from the obtained average value.

For example, in Figure 6(b) drivers with numbers s10 and s19 showed elements of dangerous driving with a percentage of 11.6% and 14%, and drivers with numbers s7 and s8 showed excellent driving results. These percentages are key to analyzing and comparing individual driving styles in the context of road safety.

Driving style analysis of the driver with number s4 in Figure 6(a): among the set of all his lines, 14.8% fell into the cells of the Kohonen map and were considered dangerous, which is a large indicator, since the average value differs from the analyzed road value. If we analyze the obtained percentage of violations (14.8%) in the time range of driving as shown in Figure 6(a), we can obtain the time interval in which the violations were detected. For example, the analyzed driver number S4 drove for an hour, of which 9 minutes he exhibited a dangerous driving style.

More detailed monitoring of an individual driving style with an assessment of the given characteristics should be carried out using statistical methods in Figure 1. For example, a sudden increase in speed demonstrates that driving is dangerous. Analyzing the behavior of the driver, taking into account the presented statistics, it is possible to pay attention to the transitions of the change in driving style, for example, by monitoring the increase in one or another indicator (increase in the danger of driving) or its decrease (transition from bad driving to good). Figure 7 in the table shows a fragment of statistics on vehicle operation obtained from the OBD-2 interface for driver number s11. These statistics are calculated by the following indicators: report time interval equal to one second, speed; engine load; intake manifold pressure; mass airflow; throttle position; acceleration. The data in the table allows to identify the zones of normal operation mode (marked in green color) and the zones with exceeded threshold value, in which dangerous driving style is formed (marked in red color). For example, in line 48, a driver numbered s11 may have a dangerous driving style identified by a combination of indicators such as over acceleration and the presence of a sharp turn. Based on the results of these statistics, it is possible to evaluate driving style both by individual indicators and by their combinations.

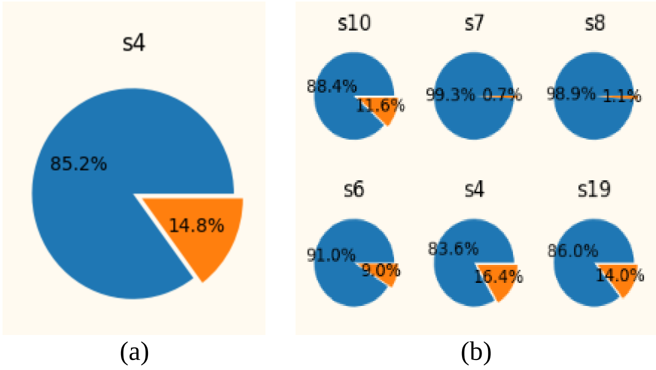


Figure 6. Pie charts, to automate monitoring of both individual drivers (a) and the group (b) with assessment of elements of dangerous driving

Data

	Driver	Speed	MAF	Engine load	Throttle Pos	IMF	Bearing	Acceleration	Color
43	s11	29.0	14.84	98.0	32.9	59.0	-2.9040625295876623	10.0	red
44	s11	34.0	14.12	63.9	31.0	35.0	-1.1475641443531401	5.0	green
45	s11	31.0	10.81	43.1	29.4	20.0	-0.03265692159624223	3.0	green
46	s11	18.0	3.37	19.2	19.2	70.0	-0.2594952562016317	13.0	green
47	s11	28.0	17.64	100.0	51.8	85.0	2.2320760166632283	10.0	red
48	s11	37.0	22.55	100.0	36.1	78.0	1.4847858169783876	9.0	red
49	s11	45.0	27.65	93.3	36.5	77.0	0.12017510484167815	8.0	red
50	s11	48.0	31.91	92.9	37.3	53.0	-0.07156585885616096	3.0	red
51	s11	49.0	22.33	63.1	34.1	35.0	-0.06661215211386207	1.0	green

Figure 7. Effects of selecting different switching under dynamic condition

The assessment of driving dynamics also allows us to identify transitions from dangerous to safe driving style. Figure 8 shows the transition from a safe driving style (first row of the table, speed 69 km/h) with increasing speed to a speed of 74 km/h (second row), which was recognized as dangerous driving. After 3 seconds (first column of the table) the driver reduced the load on the engine and returned to the previous speed. As a result, driving became safe again.

2632	s2	69.0	14.26	51.0	31.0	50.0	0.022123688451557655	0.0	green
2633	s2	74.0	29.08	100.0	36.9	40.0	0.006166439046666028	5.0	red
2634	s2	76.0	21.52	70.2	33.7	29.0	-0.0510293169991769	2.0	red
2635	s2	75.0	11.6	36.5	29.8	35.0	-0.01065755672507862	1.0	red
2636	s2	68.0	10.15	43.1	28.2	19.0	-0.01923722525231142	7.0	green

Figure 8. Table with rows of lines defining the transition to the safe style

However, a dangerous driving style can be characterized not only by speeding, overtaking or sudden acceleration, but also by sudden braking. Figure 9 shows a driver at an average speed accelerated, changed direction and reduced speed (second and third rows of the table) under the influence of a change in the surrounding traffic situation. This maneuver can be considered a dangerous maneuver. In this situation, it can be assumed that there was an overtaking of a sharply braked vehicle ahead.

The results of the experiment in Figures 5 to 9 showed that it is possible to identify the features of vehicle control both for a group of road users and for individual drivers. The developed set of quantitative characteristics for evaluating the vehicle functioning allows to construct threshold values, exceeding of which can be identified as a possible abnormal situation as shown in Figures 7 to 9. Numerical methods were used to estimate the driving style identification error. The average distance of each input vector to its BMU was calculated according to (1). Figure 10 shows the dependence of the decrease in the fall in error with an increase in training iterations.

3991	s4	24.0	4.53	42.0	21.6	74.0	0.10325996598504616	2.0	red
3992	s4	40.0	18.26	88.6	33.3	71.0	0.8800480562440498	16.0	red
3993	s4	51.0	29.79	89.8	36.5	20.0	-0.11643130975056692	11.0	red
3994	s4	38.0	4.09	20.4	20.8	32.0	0.06990507775009291	13.0	red

Figure 9. The table with row data identifying changes in direction and speed reduction

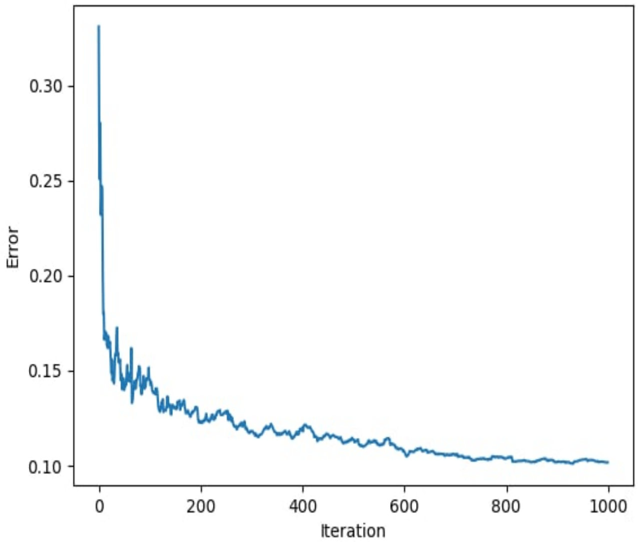


Figure 10. Graph of the dependence of error reduction with increasing training iterations

It can be seen that this parameter, initially equal to 0.35, decreases rather quickly, which indicates the correct choice of weights and the achievement of an acceptable value with an error of 0.1. Similar error estimation results for the Kohonen network are presented in [21], but the method in [21] does not take into account the dynamics of environmental changes in contrast to our approach. Thus, the proposed approach takes into account real-world scenarios of the road situation and provides a more accurate assessment. Compared to the results obtained by other researchers, the approach we developed allows to make decompositions of the feature space of driver behavior characteristics depending on the environment. The developed set of parameters about the state of the road environment and characteristics of vehicle functioning (the longitude-latitude pair in which the car is currently located; speed; engine load; intake manifold pressure; mass airflow; throttle position; acceleration) makes it possible to evaluate the influence of the dynamics of environmental changes on the dynamics of vehicle control. The obtained estimates may allow us to switch to automated determination of driving style and monitoring of vehicle behavior in order to optimize their management.

The presented method provides the basis for the development of more complex models for assessing the behavior of road users and their classification by types of dangerous driving. To further improve this approach, it is proposed to expand the set of indicators that identify dangerous driving. It is also advisable to develop scenarios of driver behavior when a particular danger is detected and automate the process of recognizing dangerous driving for individual groups of drivers. These steps can significantly improve the efficiency of the system and its ability to predict and respond to various situations on the road.

In the future, it is also planned to extend the method to analyze the behavior of drivers on individual road sections, taking into account the time intervals of traffic and its delays. This will make it possible to identify sections dangerous for road traffic, sections where traffic delays (traffic jams) occur on a permanent basis. Information about such road sections will allow to dynamically adjust the traffic situation depending on the complexity of the travel route.

5. CONCLUSION

It is found that there is a growing trend of using statistical analysis techniques to create a dynamic vehicle profile and neural network techniques to determine driving style. But there is a problem in the solutions proposed by various groups of researchers to determine the influence of the environment on the driver's behavior style. To solve this problem, we have developed an approach based on a combination of image processing algorithms that determine the dynamics of environmental changes and a Kohonen network that reflects driving style by characterizing changes in vehicle functioning. This approach was tested on a dataset containing 20 descriptive features describing the functioning of the vehicle, taken from the publicly available Kaggle website. Analysis of the test results showed that it is possible to identify the driving characteristics of a vehicle for both a group of road users and individual drivers. The formed sets of quantitative characteristics for assessing the functioning of the vehicle all

The developed method can be used not only in driving style detection but also in other areas. For example, it can serve as a basis for disease identification based on a set of input feature space of a particular pathology. This approach can significantly improve the processes of diagnosis and health monitoring, providing new tools for more accurate and early detection of diseases.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement № 075-15-2022-291 on the provision of a grant in the form of subsidies from the federal budget for the implementation of state support for the establishment and development of the world-class scientific center (Pavlov Center-Integrative physiology for medicine, high-tech healthcare, and stress-resilience technologies).

REFERENCES

[1] United Nations, “With 1.3 million annual road deaths, UN wants to halve number by 2030,” *UN News*, 2021. https://news.un.org/en/story/2021/12/1107152 (accessed Jul. 31, 2023).

[2] M. Q. Khan and S. Lee, “A comprehensive survey of driving monitoring and assistance systems,” *Sensors*, vol. 19, no. 11, Jun. 2019, doi: 10.3390/s19112574.

[3] W. Wang, J. Xi, and D. Zhao, “Driving style analysis using primitive driving patterns with Bayesian nonparametric approaches,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 2986–2998, Aug. 2019, doi: 10.1109/TITS.2018.2870525.

[4] W. Han, W. Wang, X. Li, and J. Xi, “Statistical-based approach for driving style recognition using Bayesian probability with kernel density estimation,” *IET Intelligent Transport Systems*, vol. 13, no. 1, pp. 22–30, Jan. 2019, doi: 10.1049/iet-its.2017.0379.

[5] Q. Xue, K. Wang, J. J. Lu, and Y. Liu, “Rapid driving style recognition in car-following using machine learning and vehicle trajectory data,” *Journal of Advanced Transportation*, vol. 2019, pp. 1–11, Jan. 2019, doi: 10.1155/2019/9085238.

[6] T. P. Barbosa, L. A. R. da Silva, F. J. P. Pujatti, and J. C. H. Gutiérrez, “Hydraulic hybrid passenger vehicle: Fuel savings possibilities,” *Mechanics Based Design of Structures and Machines*, vol. 50, no. 1, pp. 135–153, Jan. 2022, doi: 10.1080/15397734.2020.1714447.

[7] L. Oliveira, K. Proctor, C. G. Burns, and S. Birrell, “Driving style: How should an automated vehicle behave?,” *Information*, vol. 10, no. 6, Jun. 2019, doi: 10.3390/info10060219.

[8] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, “Social behavior for autonomous vehicles,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 50, pp. 24972–24978, Dec. 2019, doi: 10.1073/pnas.1820676116.

[9] P. Hang, C. Lv, C. Huang, J. Cai, Z. Hu, and Y. Xing, “An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14458–14469, Dec. 2020, doi: 10.1109/TVT.2020.3040398.

[10] D. Farooq *et al.*, “Analyzing the importance of driver behavior criteria related to road safety for different driving cultures,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, Mar. 2020, doi: 10.3390/ijerph17061893.

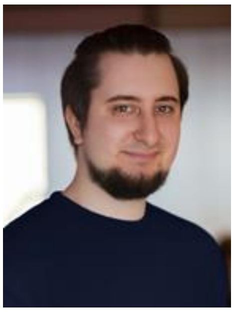
[11] J.-L. Padilla, C. Castro, P. Doncel, and O. Taubman - Ben-Ari, “Adaptation of the multidimensional driving styles inventory for

Spanish drivers: Convergent and predictive validity evidence for detecting safe and unsafe driving styles," *Accident Analysis & Prevention*, vol. 136, Mar. 2020, doi: 10.1016/j.aap.2019.105413.
[12] J. Cordero, J. Aguilar, K. Aguilar, D. Chávez, and E. Puerto, "Recognition of the driving style in vehicle drivers," *Sensors*, vol. 20, no. 9, May 2020, doi: 10.3390/s20092597.
[13] I. Silva and J. Eugenio Naranjo, "A systematic methodology to evaluate prediction models for driving style classification," *Sensors*, vol. 20, no. 6, Mar. 2020, doi: 10.3390/s20061692.
[14] G. Delnevo, P. Di Lena, S. Mirri, C. Prandi, and P. Salomoni, "On combining big data and machine learning to support eco-driving behaviors," *Journal of Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0226-z.
[15] G. Li, F. Zhu, X. Qu, B. Cheng, S. Li, and P. Green, "Driving style classification based on driving operational pictures," *IEEE Access*, vol. 7, pp. 90180–90189, 2019, doi: 10.1109/ACCESS.2019.2926494.
[16] Q. Wang, R. Zhang, Y. Wang, and S. Lv, "Machine learning-based driving style identification of truck drivers in open-pit mines," *Electronics*, vol. 9, no. 1, Dec. 2019, doi: 10.3390/electronics9010019.
[17] S. Würtz and U. Göhner, "Driving style analysis using recurrent neural networks with LSTM cells," *Journal of Advances in Information Technology*, pp. 1–9, 2020, doi: 10.12720/jait.11.1.1-9.
[18] Y. Liu, J. Wang, P. Zhao, D. Qin, and Z. Chen, "Research on classification and recognition of driving styles based on feature engineering," *IEEE Access*, vol. 7, pp. 89245–89255, 2019, doi: 10.1109/ACCESS.2019.2926593.
[19] E. Lattanzi, G. Castellucci, and V. Freschi, "Improving machine learning identification of unsafe driver behavior by means of sensor fusion," *Applied Sciences*, vol. 10, no. 18, Sep. 2020, doi: 10.3390/app10186417.
[20] M. U. Ahmed and S. Begum, "Convolutional neural network for driving maneuver identification based on inertial measurement unit (IMU) and global positioning system (GPS)," *Frontiers in Sustainable Cities*, vol. 2, Jul. 2020, doi: 10.3389/frsc.2020.00034.
[21] S. Lakshminarayanan, "Application of self-organizing maps on time series data for identifying interpretable driving manoeuvres," *European Transport Research Review*, vol. 12, no. 1, Dec. 2020, doi: 10.1186/s12544-020-00421-x.
[22] M. V. N. de Zepeda, F. Meng, J. Su, X.-J. Zeng, and Q. Wang, "Dynamic clustering analysis for driving styles identification," *Engineering Applications of Artificial Intelligence*, vol. 97, Jan. 2021, doi: 10.1016/j.engappai.2020.104096.
[23] Y. Guo, X. Wang, Y. Huang, and L. Xu, "Collaborative driving style classification method enabled by majority voting ensemble learning for enhancing classification performance," *PLOS ONE*, vol. 16, no. 7, Jul. 2021, doi: 10.1371/journal.pone.0254047.
[24] N. Jain and S. Mittal, "Bayesian Nash Equilibrium based gaming model for eco-safe driving," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 7482–7493, Oct. 2022, doi: 10.1016/j.jksuci.2021.07.004.
[25] P. Hang, C. Lv, Y. Xing, C. Huang, and Z. Hu, "Human-like decision making for autonomous driving: A noncooperative game theoretic approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2076–2087, Apr. 2021, doi: 10.1109/TITS.2020.3036984.
[26] R. Ke, Z. Li, J. Tang, Z. Pan, and Y. Wang, "Real-time traffic flow parameter estimation from UAV video based on ensemble classifier and optical flow," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 54–64, Jan. 2019, doi: 10.1109/TITS.2018.2797697.
[27] Z. Chen, R. Khemmar, B. Decoux, A. Atahouet, and J.-Y. Ertaud, "Real time object detection, tracking, and distance and motion estimation based on deep learning: Application to smart mobility," in *2019 Eighth International Conference on Emerging Security Technologies (EST)*, Jul. 2019, pp. 1–6, doi: 10.1109/EST.2019.8806222.
[28] O. Törő, T. Bécsi, and P. Gáspár, "PHD filter for object tracking in road traffic applications considering varying detectability," *Sensors*, vol. 21, no. 2, Jan. 2021, doi: 10.3390/s21020472.
[29] M. Dimitrievski, P. Veelaert, and W. Philips, "Behavioral pedestrian tracking using a camera and LiDAR sensors on a moving vehicle," *Sensors*, vol. 19, no. 2, Jan. 2019, doi: 10.3390/s19020391.
[30] T. Kohonen, "Self-Organizing Maps", *Berlin-New York: Springer-Verlag*. First edition 1989, second edition 1997, extended edition 2001.
[31] C. Gatsonis, S. C. Morton, "Methods in comparative effectiveness research," 1st Edition, *Chapman and Hall/CRC*, 2021, doi: 10.1201/9781315159409
[32] LMBC Campos, "Evaluation of Integrals of Gaussian Functions," *Generalized Calculus with Applications to Matter and Forces, Boca Raton: CRC Press*, 2014.
[33] C. Barreto, "OBD-II datasets," *Kaggle*. Accessed: Jul. 31, 2023. [Online]. Available: https://www.kaggle.com/cephasax/obdii-ds3
[34] G. Vettigli, "MiniSom," *GitHub*. Accessed: Jul. 31, 2023. [Online]. Available: https://github.com/JustGlowing/minisom

BIOGRAPHIES OF AUTHORS

Yulia Shichkina doctor of technical sciences, head of the Department Technologies of Artificial Intelligence in Physiology and Medicine as part of the International Innovative Institute of Artificial Intelligence, Cybersecurity and Communications named after V.I. A.S. Popov of the St. Petersburg State Electrotechnical University "LETI". Her research interests include mathematical modeling, parallel computing on MPP and SMP systems, optimization of access to databases, application of AI methods in applied problems. Under her leadership, international projects have been and are currently being implemented: USA "Safety applications enabled by internet-of-things infrastructure and big data analytics", Cuba "Creation of information and technological support for the study of Parkinson's disease, taking into account the collection and processing of large-scale data in real time" and Russian projects: "Creation of domestic high-tech production of vehicle security systems based on a control unit and intelligent sensors, including millimeter radars in the 76-77 GHz range." Over the past 5 years, 2 monographs, more than 60 articles and abstracts have been published, 14 certificates of state registration of computer programs have been received. She can be contacted at email: strange.y@mail.ru.

Roza Fatkieva associate professor, Department of Computer Science and Engineering, Saint Petersburg Electrotechnical University LETI St. Petersburg. Research interests: information security, modeling of information systems. Over the past 5 years, more than 30 articles and abstracts have been published, 3 certificates of state registration of computer programs have been obtained. She can be contacted at email: rikki2@yandex.ru.



Maxim Kopylov graduated from St. Petersburg State Electrotechnical University "LETI" in 2021 with a degree in computer engineering. Currently, he works in the field of image recognition, mathematical modeling, is a developer of hardware and software complexes, a participant in several projects related to signal processing. He can be contacted at email: maksim.kopylov@gmail.com.