# Intelligent information systems

UDC 004.7:004.89 **doi: 10.20998/2522-9052.2019.3.11**

G. Titov, S. Rozsokha

National Aviation University "Kharkiv Aviation Institute", Kharkiv, Ukraine

## RECOGNITION OF ROAD OBJECTS AS A MEANS OF SOLVING THE PROBLEM OF INATTENTION ON THE ROAD

**Abstract.** The **subject** of the article is the problem of inattention on the road and its root cause. The **topic** of this article is methods for recognizing objects on the road in real-time. The **tasks** of this article are to investigate the statistics of the causes of road accidents, to analyze the capabilities of modern methods of object recognition and computer vision, and to analyze the possibility of integrating these methods and models into the smartphone application. The end **goal** of research and analysis is to try to develop software that can help visually impaired and inattentive drivers, as well as reduce the potential number of road accidents. In this article used next methods: contour analysis methods with Fourier transforms and convolutional neural network design and training methods. The results of the analysis of road accidents causes showed that the main problem on the road - is the inattention of drivers. The analysis of models, methods and possibilities of their integration showed that modern methods of object recognition are capable of solving the task, but there are many factors that affect the accuracy of these methods. Recognition accuracy is also highly dependent on the smartphone and its processing power, camera, etc. **Conclusion**. Achieving the goal is possible, but modern algorithms and smartphones can solve it only partially. Also, we do not really know if it potential developed app will help driver and would not confuse them more on the road.

**Keywords:** recognition; driver; inattention; models; neural networks; technology; cars; smartphone.

### Introduction

Today in the world on average every 3.5 minutes there is a road accident [1]. The main cause of the road accidents is the neglect of the drivers [2]. Now there are many corporations and techno companies trying to solve this problem. They create different systems with “eyes” for recognizing situations and objects on the road [3]. But these solutions are very expensive and can only be installed in certain cars. To solve the problem of inattention on the road, you need a solution (technology, device, software) that would satisfy the conditions of easy installation and relative cheapness. The key to solving this problem is machine learning and computer vision. Today, a great number of companies use computer vision and machine learning in their products and technologies. Machine learning is now not just an idea and algorithms, but actually existing objects, models and methods. Elements and models of computer vision can replace many of the sensors and systems used in similar solutions before. So, solving the initial problem using computer vision and machine learning is a practical and relevant task. However, it should be understood that existing technologies for the development of models of artificial intelligence and machine learning are far from ideal. Therefore, this work is a study of the possibility of solving the problem. Despite all above, the problem remains relevant and a thorough study will in any case improve the methods of its solution and increase the amount of attention it receives.

### 1. Causes of the road accidents

According to statistics, which, by the way, is quite disappointing in Ukraine, almost all drivers make the same mistakes [4]. Consider the main causes of road accidents:

1. Unjustifiably risky behavior and conscious violation of traffic rules. Probably the most common cause of an accident. These include ignoring traffic signs and traffic signals, changing lanes without turning signal on, etc.

2. Driving drunk. Unfortunately, driving while intoxicated is almost the norm. Drunk driving itself has the most severe consequences. Due to blood alcohol, the driver is prone to losing control over his or her actions. In addition, in most cases, there is an unjustified self-confidence, slows down the reaction, dulls the attention and impairs the ability to analyze the situation around.

3. Over speed. It is necessary to choose the speed depending on the road conditions (in particular, the type, condition and slope of the road), visibility and technical condition of the car, etc.

4. Inattentive driving. This category includes the inattentive attitude of drivers to each other, including pedestrians, and not full traffic monitoring.

5. Inadequate behavior in extreme situations. The greatest danger here is the drivers-beginners or those who very rarely go on the road, because the ability to avoid accidents comes only with experience.

6. Non-observance of the distance. Such a violation may result in the vehicle being hit by a moving vehicle. In settlements, it is necessary to observe the distance at the rate of 0.5 m per 1 km/h speed, outside settlements 1 m per 1 km / h.

7. Violation of overtaking rules. Overtaking is one of the most responsible and frequent maneuvers associated with increased speed and limited visibility.

The statistics of the causes of accidents occurring as a percentage are shown in Fig. 1 [5].

Therefore, it can be concluded that most road accidents occur due to the disregard of signs and traffic lights because of drivers’ inattention.

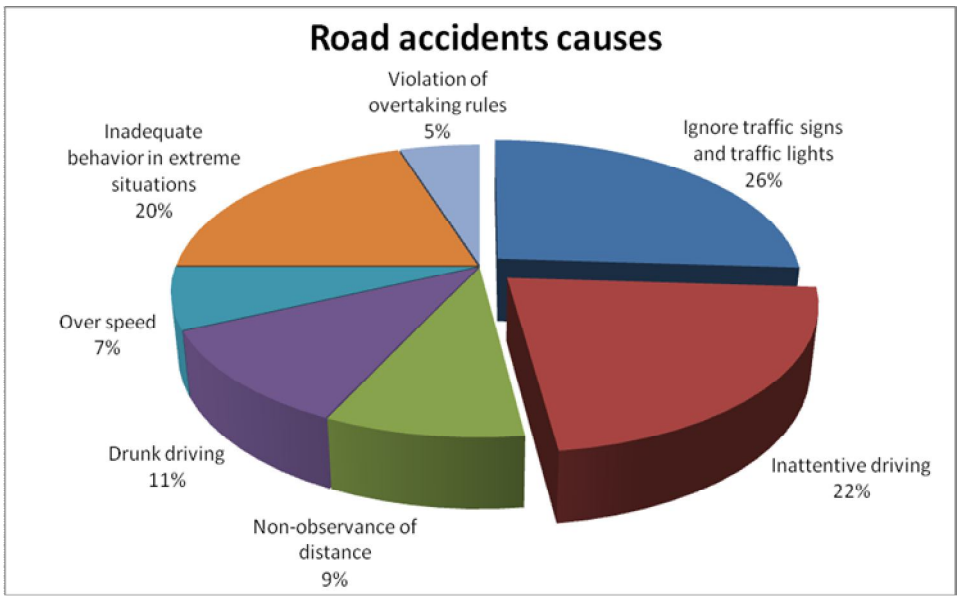


**Fig. 1.** Statistics of the causes of road accidents in percentage

Driver neglect is not always a lack of skill or habit. It can also be due to illness. In both cases, this can lead to an accident. To improve the situation, companies use specially designed tools that evaluate events on the road with cameras (or sensors) and give specific guidance to the driver of the vehicle. In some flagship car models, such systems are built-in by default. And they handle their function quite well. However, these are only some brands of high class cars. So let's understand what these tools are, what are the drawbacks of methods and algorithms. Let's find out if we can make these tools cheaper and easier than they are.

## 2. Practical part

Over the last 10 years, phones have made a qualitative leap. Before, they were only for communication, with a very limited browser and almost no video playback. But now, these are full multimedia devices with normal browsers, MS Office suite, games, equipped with cameras shooting 4K videos. Imagine, that you have complete PC in your pocket. Manufacturers are actively developing the same view [6]. From year to year, smartphone cameras are becoming more successful and more versatile. Technical innovations that are now being integrated into smartphones:

– Larger sensors. The fact is, the larger the image sensor in the camera, the better the end result

– More megapixels. The image resolution is increased.

– Wider aperture. This allows the camera to capture as much light as possible during dusk shooting.

– Optical stabilization. Optical image optimization keeps the shaking videos stable

– Improved processing. Finally, some of the achievements in smartphone shooting have been achieved not through better optics but through better processing of software made possible by more powerful phones [7].

One possible means of increasing driver inattention may be a special warning and alarm system. The main part of this system is recognition of the situation on the road and road objects, including road signs and traffic lights. Today exist and implemented the following methods of recognition of road signs:

1. Methods of contour analysis combined with Fourier transforms.

2. Methods of machine learning and neural networks [8].

Let's look at them in more detail.

**2.1 Sign recognition using contour analysis and Fourier transforms.** The OpenCV library has the necessary mathematical apparatus and a fairly high performance [9].

The steps of the task:

1. Get an image from the camera.

An example of the resulting image is shown in Fig. 2 [12].

2. Convert the resulting image.

As a result of some experiments, a filter and the following image was obtained. The result of image gray scale filtering is shown below in Fig. 3 [12].

3. Application of threshold filter [10].

The parameters for this filter were experimentally selected. The use of a pore filter is shown below in Fig. 4 [12].

4. Determination of boundaries of contours.

An optimized function from the cvCanny library, a border detector, was taken. The definition of the contours in the camera image is shown below in Fig. 5 [12].

5. Contour analysis

Any image is a two-dimensional signal. To obtain the image spectrum, one must take a one-dimensional Fourier transforms from each row of the image, and then do the same with each column of the data obtained. FFT is required for this work, since simple DFT processes the image too slowly, and should only be used for one-dimensional signals not exceeding 32 points.

Here are some properties of the Fourier transform that we need to interpret the spectrum of images.

**Fig. 2**. Sample camera image

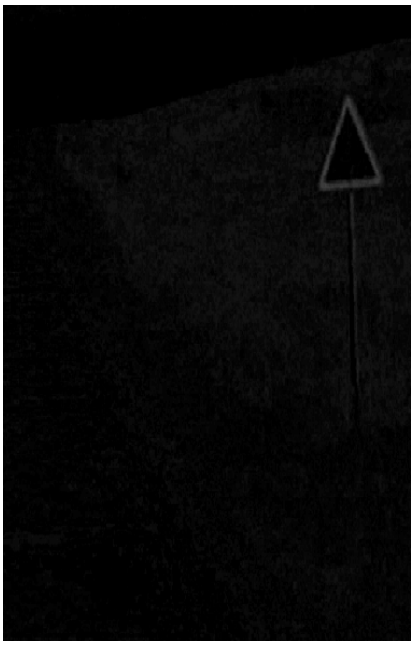

**Fig. 3.** Filter used

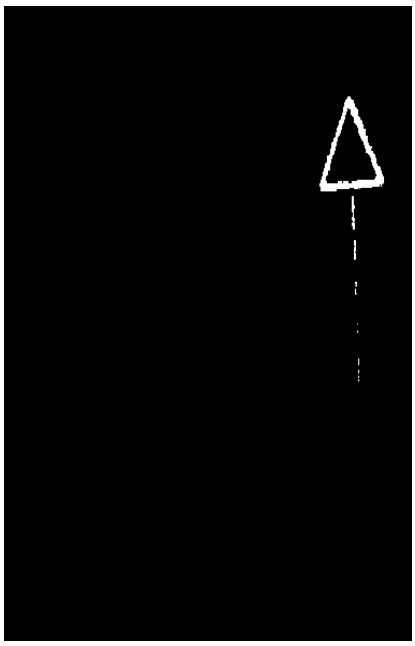

**Fig. 4**. Threshold filter

**Fig. 5**. Defined thresholds on the image

5.1. Linearity. If you add or subtract two signals in the time domain, then their spectra also add up or subtract. This means that you can spread the image across the spectrum.

5.2. Time shift. This means that when the signal is shifted in the time domain, its spectrum does not change, only the phase changes.

5.3. Rotate. This means that when the signal is rotated in the time domain, the signal in the frequency domain returns exactly the same.

5.4. Scalability. This means that as the signal expands in time, its spectrum becomes the other way around.

In the signal spectrum, the low frequency components show those parts of the image where the brightness remains virtually unchanged. High-frequency components show changes in intensity. Thus, if only the low frequencies (low pass filter) are left, then only blurry spots will be obtained after the Fourier transform. If you leave only the high frequencies (high-pass filter), then basically only the contours of the objects will remain, since it is on their borders that the brightness of the image changes dramatically. According to the Fourier property, "time shift" when changing the position of an object in the image, only the phase changes in the frequency domain, so that in amplitude it can be understood that it is this object, even if it is elsewhere. in the picture. If there are many different objects in the image, then their spectra will add up, and it will be impossible to know whether the desired object was present in the image. Of course, according to the linearity property, it is possible to subtract the spectrum of the desired object from the image spectrum. Then, if an object is present in the image, it will disappear from the image. If there was no desired object in the image, the result will depend on how many frequencies the subtracted object occupies. Namely, there will be small interferences in the image, which will make the image illegible [11].

Advantages of the contour analysis method:

– ease of implementation;
– speed of work;
– low computational complexity of algorithms.

The disadvantages of the contour analysis method:

– it is necessary to ensure that there is no noise on the image by preprocessing;
– if the object has no clear boundaries or the area is heterogeneous and contains smooth gradients, then the algorithm will not solve the problem of segmentation correctly, which will lead to the impossibility of further automated analysis;
– under different light and weather a threshold filter may need to be re-selected;
– consistent implementation;
– difficulty in finding and processing internal contours [12].

**2.2 Recognition using CNN (Convolutional Neural Networks).**

1. Visualization and intelligence analysis of data.

GTSRB is a road sign recognition datasheet in Germany. The task is to teach the classifier of road signs using marked data from GTSRB [13]. In general, the best way to get an idea of the available data is to build a histogram of train distribution and validation of datasets.

The data display is shown in Fig. 6 [14] below.

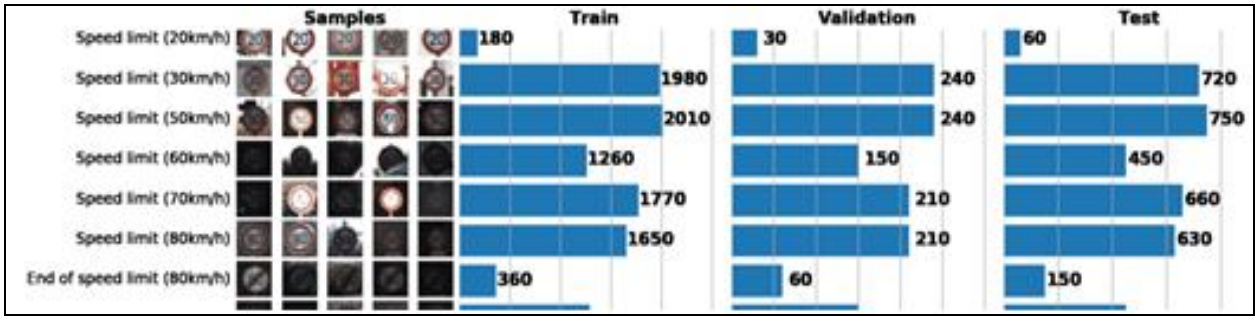


**Fig. 6**. Train, validation and test sample data

2. Normalization of images.

In order to improve the convergence of the neural network, it is necessary to bring all the images to a single illumination by converting their color gamut to grayscale. Normalization of image contrast will be accomplished using adaptive histogram normalization (CLAHE). An example of the images obtained is shown below in Fig. 7[14].

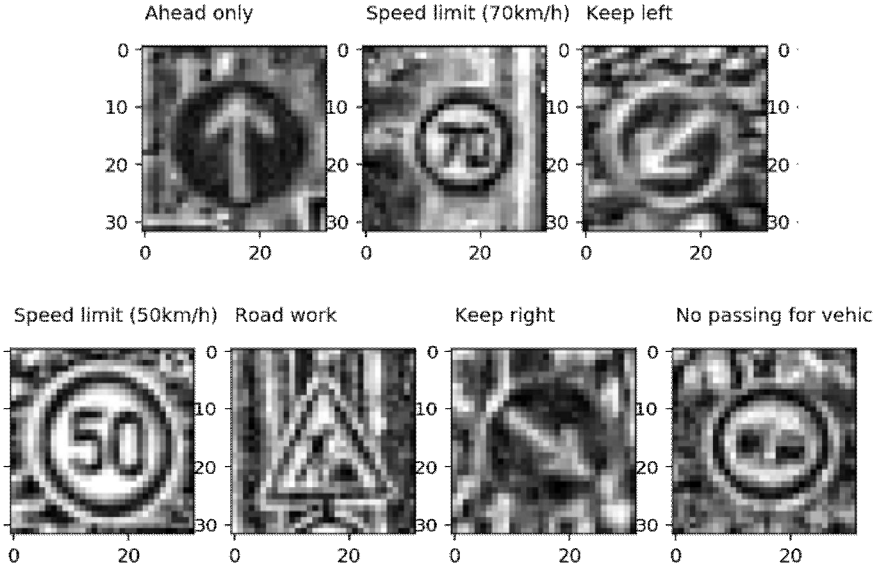


**Fig. 7.** Normalize images in grayscale

3. Data augmentation

We will construct artificial images by transforming existing images through rotation, mirroring and affine transformations. We will create new images "on the fly, so that data augmentation parameters can be quickly adjusted. Combining the network date into one large train sample, we submit it to the resulting generator. We now have two types of augmentation: the contrast and the affine transformations that apply to the samples. The generated images by augmentation are shown below in Fig. 8 [14]. Note: augmentation is required for the train set. The test set is also preprocessed, but not augmented [15].

4. Learning the neural network

After the pre-processing is complete, all the generators are ready and the set date is ready for analysis, you can go to training.

Advantages of using coagulating neural networks:

– the use of convolution kernels gives, compared to a fully connected neural network, a smaller number of tunable weights, which results in reduced time and computational resources for network learning;

– the use of convolution cores also helps to avoid pixel memory, pushing the network to summarize the displayed information;

– partial invariance of scale;

– learns using the classic error back propagation method,

– other networking and deep learning techniques may be involved;

Disadvantages of using coagulating neural networks:

– long learning time (several days or more) for a neural network with a number of convolution layers greater than two;

– the need for a large number of examples for training;

– high probability of retraining the network with insufficient number of examples;

– too many variable network parameters: the number of layers, the size of the convolution kernel for each layer, the number of nuclei for each of the layers, the shift step, the need to use sublayer layers, the degree of diminution of their dimensions, these parameters significantly affect the result, but are chosen empirically for each a new problem [16].

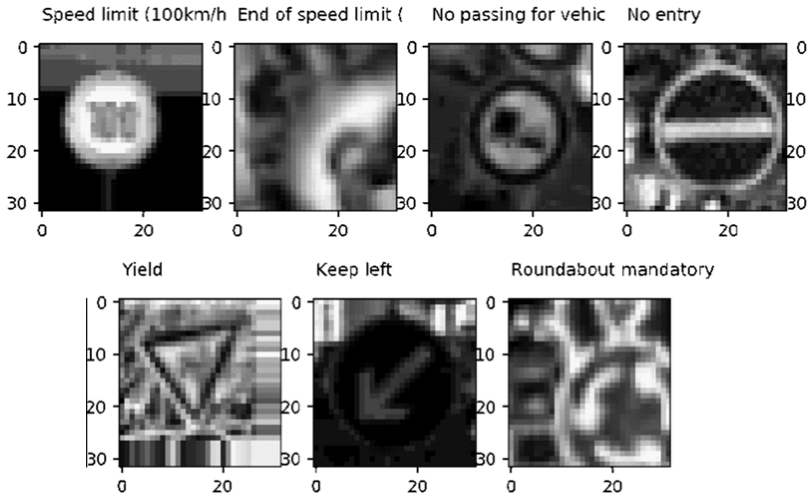


**Fig. 8**. Generated images by augmentation

## 3. Ability to integrate into the mobile application

All recognized recognition methods have several disadvantages:

1. Difficulty of use and installation.
2. The complexity of data collection and processing.
3. Complexity of implementation.
4. Great response time.
5. Not always great recognition accuracy.

Analyzing the methods, we can conclude that the basis of all systems is a powerful processor and camera. Almost all of us carry our own smartphone. Then, why can't all these methods be integrated into one mobile application? In fact, it is possible, but even if we assume that most users (or people who need this app) have an average Android smartphone, there may be some difficulties using the above libraries:

1. Different smartphones may have different cameras. It follows that the images of them may also be different.
2. Different types of color rendering also give different images in the analysis.
3. Different screen resolution, gives different-scaled images, which degrades the final performance.
4. Different performance indicators on different smartphones, also under different conditions of use of smartphones.
5. Despite the development of smartphones, not all of them have the same computing power as computers [17].

## Conclusion

The most common cause of an accident is driver inattention. Any driver in certain circumstances may be inattentive. Existing driver assistance systems are inaccessible and difficult to install. These systems are based on mathematical algorithms, methods and models of computer vision. All methods are different and have different advantages and disadvantages. Each of them solves the problem to some extent. You need an innovative tool that is easy to integrate, cheap to implement, and combines all the benefits of existing ones.

REFERENCES

1. Number of cars (2019), available at: https://www.zr.ru/content/news/350201-kolichestvo_avtomobilej_v_mire_perevalilo_za_milliard/
2. Road accidents in Ukraine (2019), available at: https://24tv.ua/ru/skolko_dtp_proizoshlo_v_ukraine_za_2018_god_oficialnye_dannye_n1094740
3. Traffic detection systems (2019), available at: http://systemsauto.ru/active/traffic_sign_recognition.html
4. Polyakov D.N. (2011), "The concept and causes of road traffic accidents", *Vestnik Omskoy yuridicheskoy akademii*, No. 3 (16), pp. 33–35.
5. Statistics of the road accidents (2019), available at: http://www.autourist.ru/info/prichini-dtp/
6. Morozov E. (2017), Smartphones power is increasing, available at: https://www.iguides.ru/main/other/sravnenie_proizvoditelnosti_sovremennykh_smartfonov_i_kompyuterov/
7. Smartphones cameras are reaching a new level (2019), available at: http://pro-spo.ru/mobilnye-texnologii-i-telefony/5338-kamery-smartfonov-proshloe-nastoyashhee-i-budushhee
8. Bychkov S.S. (2017), "Classification road sign recognition by video sequence methods", *Reshetnev readings*, Vol. 1, pp. 313–314.

9. OpenCV (2019), available at: https://opencv.org/
10. Kirsanov M.N (2015), Modification and analysis of filters for selecting contours of the image, *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota im. Admirala S.O. Makarova*, Vol. 1, pp. 122–127.
11. Dudin M.V., Povalyaev A.D., Podvalniy E.S. and Tomakova R.A. (2014), "Methods and algorithms for contour analysis for the classification of complexly structured images", *Bulletin of Voronezh State Technical University*, Vol. 10, 6, pp. 103–109.
12. Contour analysis for character recognition (2009), available at: https://habr.com/ru/post/61048/
13. GTSRB, available at: http://benchmark.ini.rub.de/?section=gtsrb&subsection=news
14. Traffic Sign Recognition Using CNN (2017), available at: https://habr.com/ru/company/newprolab/blog/334618/
15. Data Augmentation (2019), available at: https://habr.com/ru/company/smartengines/blog/264677/
16. Disadvantages of CNN (2019), available at: https://studwood.ru/1607211/informatika/dostoinstva_nedostatki_svyortochnyh_neyronnyh_setey
17. Disadvantages of modern smartphones (2015), available at: https://itc.ua/articles/preimushhestva-nedostatki-sovremennyih-portativnyih-gadzhetov/

Received (надійшла) 12.08.2019
Accepted for publication (прийнята до друку) 11.09.2019

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

**Тітов Георгій Олександрович** – студент кафедри інженерії програмного забезпечення, Національний аерокосмічний університет імені М.Є. Жуковського «ХАІ», Харків, Україна;
**Georgiy Titov** – student of Software Engineering Department, National Aerospace University "KhAI", Kharkiv, Ukraine;
e-mail: goshacrashtitov@gmail.com; ORCID ID: https://orcid.org/0000-0002-4865-9387

**Розсоха Сергій Володимирович** – кандидат технічних наук кафедри інженерії програмного забезпечення, Національний аерокосмічний університет імені М.Є. Жуковського «ХАІ», Харків, Україна;
**Sergii Rozsokha –** Candidate of Technical Sciences, Lector of the Software Engineering Department, National Aerospace University "KhAI", Kharkiv, Ukraine;
e-mail: sergey.rossokha@gmail.com; ORCID ID: https://orcid.org/0000-0002-7181-6141

**Розпізнавання об'єктів дорожнього руху як засіб рішення проблеми неуважності на дорозі**

Г. О. Тітов, С. В. Розсоха

**Анотація. Предмет** статті – проблема неуважності на дорозі, і вона є першопричиною. **Тема** цієї статті - методи розпізнавання об'єктів на дорозі в режимі реального часу. **Завдання** цієї статті - дослідити статистику причин дорожньо-транспортних пригод, проаналізувати можливості сучасних методів розпізнавання об'єктів та комп'ютерного зору та проаналізувати можливість інтеграції цих методів та моделей у смартфон. Кінцева **мета** досліджень та аналізу - спробувати розробити програмне забезпечення, яке може допомогти людям із вадами зору та неуважним водіям, а також зменшити потенційну кількість дорожньо-транспортних пригод. У цій статті використані **методи**: метод контурного аналізу з перетвореннями Фур'є та метод дизайну та навчання навчальних нейронних мереж. **Результати** аналізу причин дорожньо-транспортних пригод показали, що основна проблема на дорозі - це неуважність водіїв. Аналіз моделей, методів та можливостей їх інтеграції показав, що сучасні методи розпізнавання об'єктів здатні вирішити задачу, але існує багато факторів, які впливають на точність цих методів. Точність розпізнавання також сильно залежить від смартфона та його потужності обробки, камери тощо. **Висновок**. Досягнення мети можливо, але сучасні алгоритми та смартфони можуть вирішити її лише частково. Крім того, ми не знаємо, чи потенційно розроблений додаток допоможе водієві і не заплутає їх більше у дорозі.

**Ключові слова:** розпізнавання; водій; неуважність; моделі; нейронні мережі; технології; машини; смартфон.

**Распознавание объектов дорожного движения как средство решения проблемы невнимательности на дороге**

Г. А. Титов, С. В. Розсоха

**Аннотация. Предметом** статьи является проблема невнимательности на дороге и ее первопричина. **Тема** данной статьи - методы распознавания объектов на дороге в режиме реального времени. **Задачами** данной статьи являются исследование статистики причин дорожно-транспортных происшествий, анализ возможностей современных методов распознавания объектов и компьютерного зрения, а также анализ возможности интеграции этих методов и моделей в смартфон. Конечная **цель** исследований и анализа состоит в том, чтобы попытаться разработать программное обеспечение, которое может помочь слабовидящим и невнимательным водителям, а также уменьшить потенциальное число дорожно-транспортных происшествий. В данной статье использованы **методы**: метод контурного анализа с преобразованиями Фурье и сверточные нейронные сети, методы проектирования и обучения. **Результаты** анализа причин дорожно-транспортных происшествий, показали, что главная проблема на дороге - это невнимание водителей. Анализ моделей, методов и возможностей их интеграции показал, что современные методы распознавания объектов способны решить поставленную задачу, но существует множество факторов, влияющих на точность этих методов. Точность распознавания также сильно зависит от смартфона и его вычислительной мощности, камеры и т.д. **Вывод.** Достижение цели возможно, но современные алгоритмы и смартфоны могут решить ее лишь частично. Кроме того, мы не знаем, поможет ли потенциально разработанное приложение водителю и не будет ли оно наоборот путать их больше на дороге.

**Ключевые слова**: распознавание; водитель; невнимательность; модели; нейронные сети; технологии; автомобили; смартфон.