Image processing of transport objects using neural networks

Daniil Loktev^{1*} and *Olga Lokteva*²

¹ Bauman Moscow State Technical University (National Research University), 2-nd Baumanskaya, 5, Moscow, 105005, Russia

² Russian University of Transport (MIIT), Chasovaya str. 22/2, Moscow, 125190, Russia

Abstract. The paper is devoted to the development of an automated system model for monitoring and control of transport objects, based on the processing of images obtained using photo or video detectors, which can be installed on a fixed base near the transport highway for monitoring traffic flows and individual vehicles, and on rolling stock for monitoring transport infrastructure facilities. Image processing occurs by determining the function of blurring the image of an object, algorithms for extracting an image of an object using cascading classifiers and characteristic points, depending on the behavior of the object itself, as well as using a convolutional neural network. Machine learning of the convolutional neural network occurs when using the back propagation method of error. A neural network allows detecting objects of certain classes in the image, determining the parameters of their state and behavior. The proposed model with a movable hardware, which is responsible for obtaining the primary image, was tested on a section of the railway track to identify deviations of the state of the superstructure from the content standards, and a system with stationary photodetectors was tested to determine the parameters of moving vehicles. The obtained results of processing experimental data allowed drawing qualitative conclusions about the possibility of using the proposed algorithms and schemes for monitoring and control of various transport objects.

1 Introduction

One of the most dynamically developing areas of automation is the field of monitoring and diagnostics of transport infrastructure facilities and vehicles. Timely detection of defects, critical deformations, violations of the outer layers of the coating will largely prevent the occurrence of emergency situations or disasters [1, 2]. To configure and verify automated systems for remote control and monitoring, independent machine learning is often used, which makes information technology significantly closer to the operator and to the subject area into which these technologies are introduced, also machine learning can improve procedures and algorithms of location, capture, recognition, detection and observing moving and stationary transport objects, the appearance of which is difficult to formalize

*Corresponding author: shishkinaira@inbox.ru

and may change during their study [3, 4]. The development of schemes and procedures for machine learning and self-tuning of information-computer diagnostic and monitoring systems is largely based on the use of artificial neural networks [5, 6, 7], which in many respects are superior to other learning algorithms. The scope of neural networks is quite wide, due to the proliferation of image processing procedures for various static and moving objects in automated control, monitoring and measurement systems.

Diagnostics using automated monitoring systems is an important part of maintenance. An important factor determining the scope of the diagnostic system applicability is the possibility of placing its individual modules and blocks on ground transportation and technological complexes with different functional orientations [8, 9].

2 Research technique

To process the primary information obtained by the video monitoring system by the neural network, it is necessary to select the input network parameters that determine the state of the monitoring system, the object under study and the environment between them:

- 1) The state parameter of the system (C), which is a statistical characteristic in the form of a vector of the system parameter values at a given time t .
- 2) The state and behavior parameter of the control and monitoring system, which is a combination of changes and reactions of the monitoring system to external influences.
- 3) A parameter that defines the relationships between the individual components of the system.

Diagnostics and control of transport objects from the point of view of the complex monitoring task involves development in the following main areas:

- modernization and updating of monitoring and control tools, primarily due to integrated systems that can be placed both on rolling stock and on fixed elements near transport highways;
- automation of primary information processing from systems and means of monitoring and observation for subsequent statistical analysis and implementation of management decisions;
- formation and increase of information content of a integrated assessment of the state of transport facilities based on the analysis of the results of the work of monitoring tools and systems;
- updating of normative and technical documentation governing the organization and conduct of integrated monitoring of transport objects;
- professional retraining and advanced training of operators of mobile tools for integrated monitoring and personnel directly performing analysis and verification of the received data;
- development and implementation of innovative methods for monitoring and control of transport objects in real time.

In the optical observation range, visual-optical, visual, television, photo and video surveillance devices, night vision systems, and thermal imagers are used to obtain primary data on the behavior and condition of the studied object. Such systems seem quite promising for detecting the parameters of static and moving objects from a series of images obtained using photo and video detectors [10, 11].

In the study, it is proposed to use the method of assessing blurring of the object image and representing the image as a set of complex primitives of the Haar type or a set of corner points that uniquely describe the border of the image and background as a method of obtaining information about the object. This approach has shown its viability both in cases when the video and photographic system is moving, and the object under study is stationary, for example, when the track measuring trolley or track measuring car is

operating (Fig. 1), and in cases when the monitoring system is stationary and object under study is moving, for example, a vehicle.

Fig. 1. Diagnostic systems of the superstructure: a) track measuring trolley; b) hardware at the bottom of the track measuring car

When assessing the blurring of a stationary object (an object whose speed is significantly less than the speed of the camera) with a moving photodetector, you can use the expression

$$\Delta d = f \cdot \Delta Y / D = f \cdot v \cdot \Delta t / D = f \cdot v / (r \cdot D); \quad (1)$$

here Δd is the blurring of the object, f is the focal length of the photodetector, ΔY is the real movement of the object at a speed v during the time Δt , D is the distance from the photodetector to the object under study, r is the number of frames of the image sequence per second.

The refinement of the assessing of the image blurring or, more precisely, the border between the image of the object and the background, is possible using double synthesized blurring of the image using the function with the well-known Gaussian kernels σ_a and σ_b , this leads to two signals $b_a(x)$ and $b_b(x)$, for which the reduced difference is then calculated:

$$r(x) = (b(x) - b_a(x)) / (b_a(x) - b_b(x)) \quad (2)$$

Since the studies are devoted to the formation of a model of an information-measuring system based on the use of methods for obtaining information about an object by analyzing a series of its images, it is proposed to use convolutional neural networks allows us to combine the functionality of traditional neural networks and convolutional codes with the aim to reduce the computational complexity of the algorithms under consideration and increase the speed of data processing [12, 13].

Processing by the neural networks of the obtained primary data will determine the distance to the object under study, build a map of the image depths, detect and classify the defect of the superstructure or the condition of the moving vehicle.

The simplest neural network can be represented in the form of mathematical relations valid for an individual neuron k :

$$s_k = \sum_{j=1}^m w_{kj} x_j, y_k = \varphi(s_k + b_k) \quad (3)$$

here x_j are the input signals; w_{kj} are synaptic weights for each neuron k ; s_k is a linear combination of input actions; b_k is the threshold value for the adder; $\varphi()$ is the activation function; y_k is the output signal of a single neuron.

A feature of the work of neural networks is that the weight coefficients and threshold values that generally determine the operation of the network can be varied, this allows us to analyze different data and can be used to create neural networks with a different number of perceptron layers.

For a more accurate assessment of changes in the behavior of a neural network with small changes in weight coefficients and displacements, it is proposed to use sigmoid neurons, which apply a sigmoid function to the input data taking into account weight coefficients and threshold values:

$$\sigma = \frac{1}{1 + e^{-\sum_j w_j x_j - b}}, \text{ here } \sigma = \sigma(W * x + b) \quad (4)$$

The architecture of convolutional neural networks contains alternating convolution layers and subsampling layers, which select according to certain criteria. Using combinations of two types of layers, a convolutional neural network forms hierarchies of more complex features.

It is proposed to calculate the geometric (size, distance from the detector) and kinematic (speed, trajectory) parameters of moving and stationary objects using a convolutional neural network [14, 15, 16]. Such a network is based on the application of the convolution operation, in which each image fragment is multiplied by the convolution matrix, then summed over all elements of the matrix and written to the same position of the output image. The convolutional neural network contains alternating convolutional and subsampling layers that select elements according to certain criteria. As a result of the convolution operation, each subsequent layer emphasizes the presence of the indicated feature in the previous layer, as well as its coordinates; as a result, a feature map is generated. After the convolution operation, a subsampling operation follows, which allows us to speed up the computation operations in the neural network and make the network itself more invariant to the size of the primary image.

In the convolutional layer, the weight coefficients are training parameters, and the convolution itself can be represented as the following expression:

$$(f * g)[m, n] = \sum_{k, l} f[m - k, n - l] * g[k, l] \quad (5)$$

here f is the matrix of the original image, g is the kernel of the convolution operation.

The subsampling layers following the activation convolution layers should reduce the image dimension, there are several ways to do this, but more often than not, the maximum pooling functions are used. The operation of image convolution shows that several neural pairs can be associated with a certain weight coefficient.

An important element that ensures the operation of a neural network is its training, most often the error back propagation method is used for this procedure [17, 18, 19]. Error determination on a convolutional layer usually comes down to determining the error of a subsequent subsampling layer, if it is located in front of a fully connected layer, then the error is calculated by sequentially calculating the errors of hidden layers using the weighted error values of the output layer. If the subsample layer is in front of the convolution layer, then the back convolution operation is performed, which consists in rotating the convolution kernel 180 degrees and scanning the error map of the subsequent convolution layer with changed edge effects.

In the automated monitoring and control system under consideration, the convolutional neural network VGGNet, proposed by Simonyan and Sisserman in 2014, is used to analyze the primary images obtained by the photodetector, which suggests that a 3x3 convolution kernel can fully satisfy the computing needs of the monitoring system and will allow replacing large matrices.

As a site for testing the operation of the video monitoring system, a single-track railway section in the Moscow region was chosen [20, 21, 22]. Out of several hundred primary images obtained using the photodetector of a track measuring trolley, automatically

selected were those that contained external defects or deviations from the design position of the elements of the superstructure. In fig. Figure 2 shows the areas with an above-standard level of rubble sleeper boxes (Fig. 2a), turnout with significant ballast contamination and the formation of a plant layer (Fig. 2b).

Fig. 2. Primary images of the superstructure: a) filling of sleeper boxes with rubble, b) high contamination of the ballast

Consider the work of the proposed information and algorithmic support for a stationary monitoring and control system. We will obtain images for this system using a smartphone's camera with the following parameters: focal length $f=52$ mm, aperture value from 0.4 to 1.2 m, depth of field is in the range of 0.01-0.05 m [23, 24, 25]. A series of images were obtained for various vehicles: in Fig. 3 shows a car moving from left to right, and a bus moving in the opposite direction in the far lane.

The obtained results of experimental studies made it possible to construct graphical dependences of the probabilities of errors of the first and second kind. In Fig. 4 shows the dependences of the probability of errors of the first (Fig. 4a) and second (Fig. 4b) kind on the speed of the sought-for vehicle when using pattern recognition algorithms based on complex Haar primitives and characteristic (corner) points.

In Fig. 5 shows graphical dependences of the probability of occurrence of errors of the first (Fig. 5a) and second (Fig. 5b) kind on the speed of movement of a vehicle for low contrast (bus in Fig. 4) and high contrast (passenger car in Fig. 4). The contrast k of the vehicle relative to the background image is determined by the formula:

$$k = \frac{abs(I_{ob} - I_b) \cdot \lambda^d}{I_b}, \quad (6)$$

here I_{ob} and I_b are the average intensities of the car and background, respectively; λ is the coefficient determining the external conditions; $d = s \cdot tg \alpha$ is the minimum allowable distance at which an object can be detected, s is the maximum pixel size, α is the angular resolution of the camera [26, 27].

Fig. 3. Sequence of primary images with moving objects

Fig. 4. Dependence of the probability of errors of the first a) and second b) kind on the speed of the object

Fig. 5. Dependence of the probability of errors of the first a) and second b) kind on the speed of objects of different contrast

3 Results and discussion

According to the obtained graphical dependencies, with an increase in the speed of the object under study, algorithms based on the selection of contours by characteristic points are better suited to detect its parameters and image recognition on the image, but it can also be noted that at low speeds, algorithms based on the construction of complex classifiers of type primitives of Haar have the greatest efficiency. To detect vehicle parameters, which may have a different external color, their contrast is important compared to the background image color, therefore, when designing an automated vehicle monitoring and control system, it is desirable to create a vertical layout with a special pattern and color.

4 Conclusions

In general, the use of object recognition algorithms by their images and convolutional neural networks in a single automated control and monitoring system allows for fixed time intervals to determine not only the presence of an object in the image, but also to classify and calculate its geometric and kinematic parameters of state and behavior.

References

1. B. Glusberg, V. Korolev et al. *MATEC Web of Conferences*, **239**, 01054 (2018) DOI: 10.1051/matecconf/201823901054
2. E. Gridasova, P. Nikiforov, A. Loktev et al. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 559-569 DOI: 10.1007/978-3-030-37916-2_54
3. F. Rosenblatt Psychological review. **65(6)**. 386 (1958).
4. M. Ghayoumi *Journal of Communication and Computer*. **14(1)**, 34-38 (2017).
5. K. Kavukcuoglu et al. *Advances in neural information processing systems*. 1090-1098 (2010).
6. S.J. Pan, Q. Yang *IEEE Transactions on knowledge and data engineering*. **22(10)**, 1345-1359 (2009).
7. K. Solanki, P. Pittalia, *International Journal of Computer Applications*, **133(12)**, 20-24 (2016).
8. A. Loktev, V. Korolev et al. *IOP Conf. Series: Materials Science and Engineering*, **463**, 032019 (2018) DOI: 10.1088/1757-899X/463/3/032019
9. A. Savin, A. Kogan et al. *International Journal of Innovative Technology and Exploring Engineering*, **8(7)**, 2325-2328 (2019)
10. B. Glusberg, A. Loktev et al. *Advances in Intelligent Systems and Computing*, **982**, 337-345 (2020) DOI: 10.1007/978-3-030-19756-8_31
11. A. Lyudagovsky, A. Loktev, V. Korolev et al. *E3S Web of Conferences*, **110**, 01017 (2019) DOI: 10.1051/e3sconf/201911001017
12. S. Klyuev, A. Klyuev, N. Vatin. *MATEC Web of Conferences*. 2018. **245**. DOI:10.1051/matecconf/201824503006.
13. A. Loktev, V. Korolev et al. *Vestnik of the Railway Research Institute*, **77(2)**, 77-83 (2018) DOI: 10.21780/2223-9731-2018-77-2-77-83
14. A. Loktev, V. Korolev et al. *Procedia Engineering*, **189**, 133–137 (2017) DOI: 10.1016/j.proeng.2017.05.022
15. B. Glusberg, A. Savin et al. *Advances in Intelligent Systems and Computing*, **982**, 556-570 (2020) DOI: 10.1007/978-3-030-19756-8_53
16. B. Glusberg, A. Savin et al. *Advances in Intelligent Systems and Computing*, **982**, 571-587 (2020) DOI: 10.1007/978-3-030-19756-8_54
17. B. Kovačič, R. Kamnik, A. Pustovgar, N. Vatin. *Procedia Engineering*. 2016. **165**. Pp. 918–925. DOI:10.1016/j.proeng.2016.11.801.
18. B. Glusberg, V. Korolev et al. *E3S Web of Conferences*, **138**, 01017 (2019) DOI: 10.1051/e3sconf/201913801017
19. A.A. Loktev, V.V. Korolev, E.A. Gridasova. *IOP Conf. Series: Materials Science and Engineering*, **687**, 022036 (2019) DOI: 10.1088/1757-899X/687/2/022036
20. V. Korolev. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 621-638 DOI: 10.1007/978-3-030-37916-2_60
21. A. Savin, O. Suslov et al. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 648-654 DOI: 10.1007/978-3-030-37916-2_62
22. V. Korolev, A. Loktev et al. *IOP Conference Series: Earth and Environmental Science*, **403**, DOI: 10.1088/1755-1315/403/1/012194
23. A. Loktev, V. Korolev et al. *Vestnik of the Railway Research Institute*, **77(6)**, 331-336 (2018) DOI: 10.21780/2223-9731-2018-77-6-331-336

24. A. Savin, V. Korolev et al. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 797-808 DOI: 10.1007/978-3-030-37916-2_78
25. A.V. Savin, V.V. Korolev, I.V. Shishkina. *IOP Conf. Series: Materials Science and Engineering*, **687**, 022035 (2019) DOI: 10.1088/1757-899X/687/2/022035
26. V. Korolev. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 175-187 DOI: 10.1007/978-3-030-37919-3_17
27. I. Shishkina. *TransSiberia 2019: VIII International Scientific Siberian Transport Forum*, 834-844, DOI: 10.1007/978-3-030-37916-2_82