On the application of neural network technologies for control problems in cognitive transportation systems

O.Yu. Lukomskaya¹ and Y.A. Seliverstov²

¹ Candidate of Technical Sciences, Associate Professor, Leading Researcher, Solomenko Institute of Transport Problems of the Russian Academy of Sciences, Saint Petersburg Electrotechnical University «LETI», Saint Petersburg, Russia
² Candidate of Technical Sciences, Senior Researcher, Solomenko Institute of Transport Problems of the Russian Academy of Sciences, Saint Petersburg, Russia

E-mail: luol@mail.ru

Abstract. In the work, the functions of cognition are brought in line with the aspects of the functioning of cognitive transport systems and technologies. We consider the mathematical apparatus of cognitive technologies – neural network control systems, based on trained multilayer neural networks of direct action in the class of adaptive control systems for dynamic objects. The possibility and variants of their inclusion in the control system of the transport system are clarified. On the basis of the considered approach, a country road recognition system based on a convolutional neural network has been developed, which will automatically segment rural roads using satellite images and build a road network diagram to analyze the spatial development of transport networks and to control unmanned vehicles. The development of a neural network based on U-net was carried out in Python 3x. The training set consists of 880 images prepared by hand markup. The accuracy of the developed model when testing on prepared samples was 64%. According to the results of the study, conclusions were drawn and prospects for further functional development of the developed tools were determined.

Introduction
The problem of transport system management nowadays presumes the implementation of cutting edge and knowledge intensive technologies, for example, cognitive transport technologies.

Cognition (knowledge) is an action attributed to mental process through which internal and external neural processes are transformed, stored, regenerated and used. Thus cognition possesses many functions such as perception, attention, coding of memory, storage and reaction, making decisions, discussion, problem solving, image forming, planning and fulfilling of actions [1]. Let us attribute the functions of human mind to the functional possibilities of transport systems and technologies at the stage of designing cognitive transport systems and technologies (CTST) and look at the tasks based on this concept (table 1).

Major mathematical tools and their properties, which are used in design and modeling of control systems in CTST [2–8] include:

- artificial neural network (ANN), the prototype of which is the neural system of a living creature, and the element of which is the base processor element (BPE) – the analog of biological neuron, imitating the most important basic functions of human mind, of biological
neuron (table 1). The main property of ANN which is used in control is learning, besides there are signals of direct distribution, universal approximation properties, adaptive properties of control structures with ANN and capability of parallel processing of analog and discrete signals for multidimensional objects [2–7];

- artificial intellect, which has the possibility to form new rules and relations, actualizing the process of decision making;
- self-organization that uses the approach of the system of organization development under the influence of inner reasons and mechanisms of adaptation to the condition of environment.

| Functions of human mind (cognition) | Functions of cognition in transport technologies |
|-----------------------------------|-------------------------------------------------|
| perception                        | receiving of information (data) from detectors and sensors |
| attention                         | analysis of incoming information (organization of data from sensors) |
| memorizing storage                | coding and structuring of information (semantic processing of organized data) |
| reaction                          | structuring and recording of data |
| decision making                   | feedback (reaction) of system to ingoing effects (in the form of a dialog, sound and video signal) |
| discussion                        | development of decision making (formation of knowledge necessary for decision-making) |
| problem solving                   | context processing of information on demand |
| image forming planning            | forming and output of a vector of control actions, which aim at the implementation of changes in the state of executive mechanisms and devices |
| action fulfillment                | transmission and processing of video signal for informational exchange |
|                                   | planning (schedule forming, tough coordinate work regime forming) |
|                                   | management of actions of executive devices, which bring the physical environment into the required condition based on the vector of control actions |

**1. Description of multi-layer neural network in control system**

Multilayer neural networks (MNS) or multilayer perceptrons perform the function of an adaptive regulator in a dynamic control system. At the input of the executive device of the system with the objective function $Z = \text{min}$, the MNS in the learning process forms the optimal control action (the goal of training and the goal of object control coincide, i.e., one common $Z$ is set). However, the functioning of the network can take place in 2 stages and, accordingly, the target functionals of training the network and control of the object may differ from each other: 1) preliminary training of network follows the optimal control law, which is done on the basis of the theory in accordance with the given functional of learning; 2) displaying of optimal control function in the output of the network or in the input of the working device when the network is connected to a real system [2-7]. MNN in dynamic systems of control performs the function of adaptive regulator, identifier of condition and optimizer, and is called a neural network control system [4-7].

Classification of MNN according to architectural parameters is given in the fig 1.

In adaptive control MNN is used as an approximator of functions of several variables (as regressors with sigmoid function of activation with transmitting function). The peculiarities of regressors depend on the way of inclusion in the control system [6].

Mathematical model of object control is written as equations of state, which describe dynamic system in Cauchy academic form for adaptive control:

---

Table 1. Functions of cognition in transport systems and technologies
where $\mathbf{x}$ is the vector of state variables; $\mathbf{u}$ is the vector of controlling influence $\Theta \in \Omega \in \mathbb{R}^d$ is the vector of unknown environmental influence.

A well-known hybrid MNN – Neuro-fuzzy control system – is an MNN, which uses mechanisms of fuzzy logic conclusion (fuzzy-rules: IF … THEN) in the tasks of control (or, otherwise, MNN learning with data, which are produced by fuzzy-system). Structurally the following layers are identified: «fuzzification» – «fuzzy-rules» – «defuzzification» [2]. The practical use of above mentioned is covered in many research works, including [4–9].

![Classification of MNN](image)

Fig. 1. Classification of MNN.
2. Using multilayer convolutional neural networks to develop a country road recognition system

The overall growth of cities in the past two decades has led to a significant development of transport networks. An actively developing road infrastructure requires frequent updates of road maps. A wide range of applications depend on this information, such as intelligent transportation systems [10 (4)], monitoring the spatial development of a city [11], automatic data updates for geolocation systems, or supporting rescue operations in disaster relief. A satellite equipped with synthetic aperture radar (SAR) can scan the terrain. The resulting physical information about the terrain is more resistant to changes in exposure and color than optical images. In addition, SAR sensors can operate regardless of any weather conditions, which is a major advantage when capturing a region affected by a natural disaster.

Here are some recent research papers in this area. For the first time, deep convolutional neural networks (DCNN) demonstrated unrivaled efficiency in image analysis in 2012 [12]. While many DCNN architectures specialize in image classification (predicting one tag from an image: airplane, car, ship, etc.) [13] [14], others have achieved good performance in remote sensing tasks such as semantic tagging of aerial photographs [15] [16]. To maximize the spatial accuracy of the information extraction, we perform pixel segmentation using complete convolutional neural networks (FCNNs). Introduced in 2015, the convolutional neural network FCN8s [17] trained on the popular POSCAL VOC 2012 reference datasets set a new record for the quality of semantic segmentation [18]. In [19], a technique was proposed for the automatic segmentation of satellite images into several classes, such as buildings, rivers, roads, etc. In [20], a segmentation system for roads of various classes of roads was developed based on a cascade network and a direction module to capture the linear structure of the road.

The present study focuses on road extraction from SAR satellite imagery using deep convolutional networks.

To solve the problem of road segmentation on satellite SAR images, at the first stage, an algorithm of the system was developed (Fig. 2).

The neural network contained 23 convolutional layers (Figure 3). It consists of a tapering path (left) and an expanding path (right). The tapering path is a typical convolutional neural network architecture. It consists of reapplying two $3 \times 3$ convolutions, followed by a ReLU activation function and a maximum combining ($2 \times 2$ power of 2) operation for downsampling. Property channels are doubled at each downsampling step.

Fig. 2. Algorithm of the road segmentation system.
Fig. 3. U-net neural network diagram for road recognition.

Each step in the expanding path consists of an upsampling property map operation, followed by:
- convolution $2 \times 2$, which reduces the number of property channels;
- combining with an appropriately cropped property map from the collapsing path;
- two $3 \times 3$ convolutions followed by a ReLU (Figure 4).

Cropping is necessary due to the loss of border pixels with each convolution.

Fig. 4. U-net network diagram.

Each blue square corresponds to a multichannel property map. The number of channels is marked at the top of the square. The $x$-$y$ size is indicated at the bottom of the square. Squares in Figure 4 correspond to copies of the map of property and arrows correspond to various operations [15, 16].
Trashholding was used for data preprocessing. Augmentation was performed using an image generator programmed to shift, rotate and enlarge images. Data augmentation is necessary to train the network for the desired properties of invariance and stability when a limited number of training examples are available. The network was trained through the Google Colaboratory service. The network is trained using stochastic gradient descent based on the input images and their corresponding segmentation maps.

The dependence graph of the IoU coefficient on the number of epochs for the training and validation samples is shown in the fig. 5. The X-axis is the epoch. At each epoch, the entire training set is passed through the network and the weights are corrected based on these data.

![Fig. 5. Accuracy versus training epochs (IoU metric).](image)

The Y-axis is a metric for assessing the quality of IoU segmentation. Due to convolutions, the output image is smaller than the input signal by a constant border width. Applying pixel by pixel, the soft-max function computes the energy from the final property map along with the cross-entropy function.

The main advantages of the algorithm for the operation of neural networks are the ability for self-learning, realized on the basis of the analysis of precedents, as well as the high accuracy of the result [17]. At the second stage, a training sample was formed, consisting of a Training Set, a Test Set, and a Validation Set. For this, a set of 3600x3600 satellite images was selected and the roads were manually marked. As a result, three types of images were obtained: original images (Fig. 6), an image mask (Fig. 7), and images with a mask overlay (Fig. 8). The total volume of the training sample was 1200 images.

Further, the development of a convolutional neural network was carried out. The neural network was created in Python 3x using the TensorFlow, TensorBoard, Pandas, Numpy, Scipy, Matplotlib, Sklearn libraries. The U-net network was chosen as a neural network model [8]. The neural network consisted of 10 layers (Fig. 5). To train the neural network, a training sample was fed into the input, 75% of which went to training, and 15% to testing. The number of epochs for training was 60. The learning outcomes are presented in fig. 6. The binary cross entropy was used as a loss function; its value at the stage of training completion was 0.1012. The per-step mean Intersection-Over-Union (mIOU) was 0.6884.

The aggregate intersection mean is a common scoring metric for semantic image segmentation that first calculates the IOU for each semantic class and then calculates the average across the classes. IOU is defined as follows: IOU = true_positive / (true_positive + false_positive + false_negative). The predictions are accumulated in a weighted mixed matrix and then mIOU is calculated from this matrix.
The value of the cost function for the cross validation data was \( \text{val\_loss} = 0.1321 \). And the value of the recognition quality was \( \text{val\_mean\_iou} = 0.6884 \). Characteristics at one of the stages of neural network training (epoch 58-60) are shown in Fig. 9.

Comparisons of the original satellite image, the image from the training sample and the image obtained as a result of the neural network operation with road segmentation are presented in Fig. 10.

**Conclusion**

The system developed on the basis of a convolutional neural network for road segmentation on satellite images of SAR (Figure 11) has shown its high efficiency. The resulting terrain marking is comparable in quality to manual marking.

The practical implementation of this system as one of the components of an intelligent transport system [18] will allow automatic assessment of the existing roads in order to develop and organize new transport networks, and using as a marker backbone route network [19] for unmanned aerial monitoring of rural settlements.
Fig. 11. Road image segmentation process.

References

[1] Asaul A., Malygin I. and Komashinskiy V. 2017. The Project of Intellectual Multimodal Transport System. Transportation Research Procedia, pp. 25–30. doi: 10.1016/j.trpro.2017.01.006

[2] Seliverstov Y.A., Seliverstov S.A., Komashinskiy V.I., Tarantsev A.A., Shatalova N.V. and Grigoriev V.A. 2017. Intelligent systems preventing road traffic accidents in megalopolises in order to evaluate XX IEEE International Conference on Soft Computing and Measurements (SCM) (St. Petersburg) 2017, pp. 489-492, doi: 10.1109/SCM.2017.7970626

[3] Krizhevsky A., Sutskever I. and Hinton G. E. 2012. ImageNet Classification with Deep Convolutional Neural Networks (Proc. NIPS, Lake Tahoe), pp. 1097–1105.

[4] He K., Zhang X., Ren S., and Sun J. 2016. Deep Residual Learning for Image Recognition (Proc. CVPR, Las Vegas), pp. 770–778.

[5] Tyukin I.Yu., Prokhorov D.V., Terekhov V.A. 2003. Adaptive control with nonconvex parameterization (IEEE Transactions on Automatic Control), vol 48, no 4, pp. 554–567.

[6] Sherrah J. 2016. Fully Convolutional Networks for Dense Semantic Labelling of High-Resolution Aerial Imagery (CoRR), vol. abs/1606.0.

[7] Geng J., Wang H, Fan J. and Ma X. 2017. Deep Supervised and Contractive Neural Network for SAR Image Classification (IEEE Trans. Geosci. Remote Sens.), vol. 55, no. 4, pp. 2442–2459.

[8] Long J., Shelhamer E. and Darrell T. 2015. Fully Convolutional Networks for Semantic Segmentation (Proc. CVPR), pp. 3431–3440.

[9] Everingham M., Van Gool L., Williams C. K., Winn J. and Zisserman A. 2010. The pascal visual object classes (VOC) challenge. Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–338.

[10] Mohammed El Amin A., LiuQ., & Wang Y. 2016. Convolutional neural network features based change detection in satellite images (First International Workshop on Pattern Recognition) doi:10.1117/12.2243798

[11] Guangliang Cheng, Chongrue Wu, Qingqing Huang, Yu Meng, Jianping Shi, Jiansheng Chen, Dongmei Yan. 2019. Recognizing Road From Satellite Images by Structured Neural Net-work, Neurocomputing (In Press, Corrected Proof) DOI: doi.org/10.1016/j.neucom.2019.05.007

[12] Seliverstov S.A. et al. 2017. Elaboration of intelligent development system of megalopolis transportation. IEEE II International Conference on Control in Technical Systems (CTS) (St. Petersburg), 2017, pp. 211–215. DOI:10.1109/CTSYS.2017.8109528
[13] Seliverstov S. A. and Seliverstov Y.A. 2018. Developing principles for building transport networks of conflict-free continuous traffic. Transportation research procedia 36 pp. 689–699. DOI:10.1016/J.TRPRO.2018.12.122

[14] Skorokhodov D. et al. 2018. Using Augmented Reality Technology to Improve the Quality of Transport Services. In: Sukhomlin, V., Zubareva, E. (eds). Convergent Cognitive Information Technologies. Communications in Computer and Information Science. (Springer, Cham.) Vol. 1140, pp. 339–348. DOI:10.1007/978-3-030-37436-5_30

[15] Ronneberger O., Fischer P., Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Navab N. Hornegger J, Wells W, Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science (Springer, Cham). Vol. 9351.

[16] Malygin I. G. et al. 2020. Mobile technologies in intelligent transportation systems. In: Sukhomlin, V., Zubareva, E. (eds). Convergent Cognitive Information Technologies. Communications in Computer and Information Science. (Springer, Cham.) Vol. 1140, pp. 384–391. DOI:10.1007/978-3-030-37436-5_33

[17] Seliverstov Y. A, Seliverstov S.A., Lukomskaya O.Y., Nikitin K.V., Grigoriev V.A. and Vydrina E.O. 2017. The method of selecting a preferred route based on subjective criteria. IEEE II International Conference on Control in Technical Systems (CTS) (St. Petersburg), 2017, pp. 126–130. DOI: 10.1109/CTSYS.2017.8109506.

[18] Kripak M.N., et al. 2020. Analytical support for effective functioning of intelligent manufacturing and transport systems. IOP Conference Series: Materials Science and Engineering, Volume 709, Issue 3, pp. 033065. DOI: 10.1088/1757-899X/709/3/033065

[19] Seliverstov Ya.A., Chigur V.I., Sazanov A.M., Seliverstov S.A., Svistunova A.S. 2019. Sentiment analysis of "AUTOSTRADA.INFO/RU" users’ comments // Tr. SPIIRAN, 18:2, pp.354–389, https://doi.org/10.15622/sp.18.2.354-389