Intelligent Recognition in Automated Meters Surveying

Anton Ivaschenko, Arkadiy Krivosheev, Denis Sveshnikov, Nikita Svechkov, Tatiana Feschenko, Yuliya Tyshkovskaya, Alexandr Chuvakov
Samara State Technical University
Samara, Russia
anton.ivashenko@gmail.com

Abstract—**The paper proposes a new multi-layer solution to combine various algorithms implementing Artificial Intelligence (AI) for image recognition. Several neural networks are introduced to solve specific problems of objects identification. Additional "pre-launch matcher" is supplemented to scope out various objects and assigns them to the most corresponding AI modules. Distributed meter surveying is taken as an illustrative example of successful use. The introduced solution was implemented to process and analyze the results of electrical meters that are manually monitored by a group of patrol personnel inspectors using hand held devices. The results of development and testing show how the quality of neural network used for meter processing can be improved in practice.**

I. INTRODUCTION

Smart systems and devices that apply Artificial Intelligence (AI) in various problem domains become widespread nowadays. One of the efficient areas of application of artificial neural networks is image recognition. Modern solutions are capable of successful pattern identification in face and speech recognition, image processing, computer vision, finger print identification, etc.

Digit recognition in meter readings is an illustrative example of how AI based solution can be used in practice. The problem is actual for e.g. electrical power network inspectors that need a mobile tool to capture and process photo and video images that describe the current state of network devices, lines and systems [1, 2]. Despite the fact that modern meter for electricity and water are equipped with electronic components and communication channels for transmitting readings directly to the service provider for billing the consumer the problem of their visual monitoring and control remains challenging. Electrical supply companies continue employing the specifically trained personnel for visual control and maintenance of the counters in order to reduce illegal tapping and misrepresentation.

Nevertheless implementation of AI to practice remains yet challenging. The major problem resides in low versatility of neural networks caused by filtering property that lead to the lack of multitasking. In order to provide good identification quality the neural network is being especially trained to reduce the noise affected by surroundings. This noise can contain useful information that is lost as a result.

For example, the tasks of meter reading analysis include a) identification of display panel and b) digit recognition for its indication evaluation. Targeting primarily the second task improves the quality of neural network application, but limits the prospects of its practical use. Targeting both tasks by one neural network introduces difficult training and low efficiency.

To cover this gap there is proposed a combined two-layer solution in this paper. Several neural networks are introduced to solve specific problems of objects identification. Additional "Pre-launch Matcher" is supplemented to scope out various objects and assigns them to the most corresponding AI modules. It can be based by an extra neural network itself or implement preliminary defined rules or reasoning. More details are given below.

II. STATE OF THE ART

The task of data processing is an acute problem in the age of explosive growth of information [3]. World statistics confirms that if in the past the main problem was the lack of necessary information, now there is too much information and it continues to be generated with increasing speed. Data is becoming a strategic resource in almost all areas with intensive use of information that determines the competitiveness, level of development of science and industry.

Big data [4] has led to an explosive increase in the popularity of wider data mining methods, partly because there is much more information, and by its very nature and content it is becoming more diverse and extensive. When working with large data sets, relatively simple and straightforward statistics are no longer enough. The requirement to simultaneously account for a huge number of features in the data leads to the need to move from a simple search and statistical analysis of data to a more complex data mining.

Artificial Intelligence and Machine Learning algorithms are widely used nowadays for image analysis and pattern recognition [5 – 9]. The efforts of their developers are concerned with selection of the most suitable model of neural network, collecting the most efficient training dataset and thus improving the identification quality. Implementation of intelligent recognition in practice is often limited by the scope and specifics of the problem domain.

The following object detection models are traditionally used to find the position of objects in the image: Region Based Convolutional Neural Networks (RCNN), Single Shot MultiBox Detector (SSD), and You Only Look Once (YOLO) [10]. These models allow you to find the position of an object on image and classify it. To solve this problem they implement one of the two different approaches.

The first one splits the image into regions and classifies each region into different classes of objects (RCNN and its derivatives). The second one considers object detection as a problem of regression or classification (YOLO, SSD, etc.). According to the results of comparing the object detection models in [11, 12], models based on the YOLOv2 architecture show good results at almost the highest image processing speed. For mobile devices, processing speed is as important a model characteristic as accuracy. Therefore, the YOLOv2 model is one of the best for use in mobile devices.

The problem of electrical meters' photo surveying requires counter reading recognition. This task has to be successfully performed in various conditions, including weak light and darkening, overshadowing, obfuscation, occlusion and other failures.

Currently on the market there are quite diverse types of meters (see Fig. 1) both analog and digital. Recognition of the readings of the LCD and analog meters is different, this is due to the difference in the readings themselves, in the analog devices in most cases the readings are white on a dark background, the symbols have no gaps, and in the LCD readings are dark segmented values against a light background.

Fig. 1. Examples of analog and digital counters

The vast majority of them cannot transmit the values electronically and require photo surveying. To solve this problem there was developed a special software for hand held devices and smartphones supporting the operator to recognize the readings within the framework of the process collecting and further analyzing the level of energy consumption by the population of a particular region. In order to improve and facilitate the procedure it was suggested to supply the surveying software with an intelligent component.

Still the problem of automated collection and intelligent recognition of readings from all types of meters is concerned with the fact that meters have completely different specifics and interfaces. A neural network cannot be sequentially trained to solve several tasks that very and contradict to some extent. The symbolic data that in some cases is treated as information noise and should be reduced from consideration turns out to be critical in other circumstances. This leads to a certain difficulty: the readings on the counters can be of different fonts, the counters themselves can be of different types, and the quality of the shooting can vary significantly in each case.

To solve this problem it was proposed to develop an architecture including several neural networks and a dispatcher or "matcher", which takes all the data as input, determines the type of counter, and selects a neural network especially trained to recognize this type of readings. This dispatcher can implement either declarative logic in the form of rules or reasoning algorithms or a separate neural network itself being specifically trained to identify the counter type.

This solution gives new opportunities to expand the applicability of mobile neural networks for object recognition in difficult conditions.

III. PRE-LAUNCH AI MATCHER SOLUTION

Conventionally, the data processed by automated meters surveying is divided into two categories: stored on a digital medium, processed on demand, and delivered from all kinds of sensors that require processing in real time. Due to the different types of data generated and processed per unit time, classical architectures of the form "one task and one data type – one neural network" can no longer provide sufficient flexibility for new tasks on new data. To solve this problem there was developed two-layer system architecture for processing various types of data, see Fig. 2.

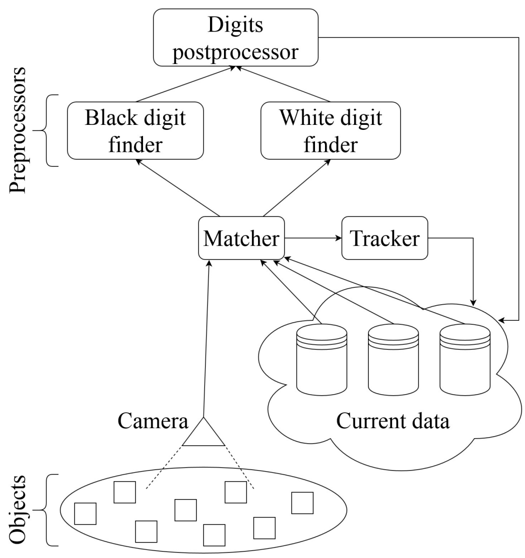

Fig. 2. Pre-launch AI matcher architecture

Camera is located on a smartphone and provides photo fixation. The data from the camera and the current data that describes the context are received and continuously sent to the Matcher. The Matcher determines the best option for data processing by choosing the modules used: two preprocessors (called Digit finders) and Tracking module. To process data, each preprocessor uses a neural network inherent to it, the results of recognition are returned for subsequent layout and transfer to the postprocessor.

The data transmitted to Digit finders includes the probability of recognizing each digit, the type of the digit, and its position in the analyzed section of the frame. The postprocessor compiles and analyzes the data received from the preprocessors and makes a decision on which result is the best to be transferred back to the Manager in the current data warehouse. Postprocessor makes attempts to find a sequence similar to a counter reading in the arrays of digits. If the Matcher selects several preprocessors for digit recognition, the array is processed in all digits, taking into account the color of each digit.

As a result, the Matcher receives combinations of digits, most likely being counter readings. The tracking module allows you to track the movement of the camera in two adjacent frames. Thus, the Matcher can compare several recognition results taken at different periods of time, and decide whether to discard the collected data as a result of the camera removal or to approve the final result according to several consecutive recognitions. Another feature of the Matcher is the ability to control camera settings, so it can focus, turn on a flashlight, adjust white balance (color balance), etc.

The Matcher can implement either deterministic rules or artificial intelligence itself. The experiments described below were carried out for a rule based Matcher. Adding a neural network will become available after the system is put into operation and allow improving the quality of automated meters surveying.

IV. Neural Network Implementation

To train the neural network, it was decided to use a simulated dataset. The reason is that a large amount of images is required, and it was not possible to collect so many photos of counters and mark them manually. The architecture of an image generator was proposed, which produces random transformations of digit symbols by scaling, darkening, blurring, rotation, and adding a background.

The Generator input contains a set of images of digits and background pictures including 3 thousand copies obtained from system fonts. A set of background images was collected manually and consists of 1 thousand copies. Each background image is initially cut into squares of a given size (an image of 416x416 pixels is fed into the tiny YOLOv2 neural network) and each one is rotated by 0, 90, 180 and 270 degrees.

Then the Generator takes a random image of a digit, and performs a series of transformations with it: scale, rotate by a small angle (up to 40 °) to the left or right, darken, etc. The modified symbol of a digit is inserted into a random place on the background image. Then the Generator performs a brightness check: the digits are inserted only where the background brightness is lower than the brightness of the number. The area in the image into which the digit is inserted is remembered and recorded in the xml markup file, together with the digit symbol. Thus, a dataset for training a neural network is formed.

To train the neural network, a synthetic data set was generated containing 4000 images with randomly located digits. The resulting image was also likely to be blurred. Examples of the resulting images for training a neural network are shown in Fig. 3.

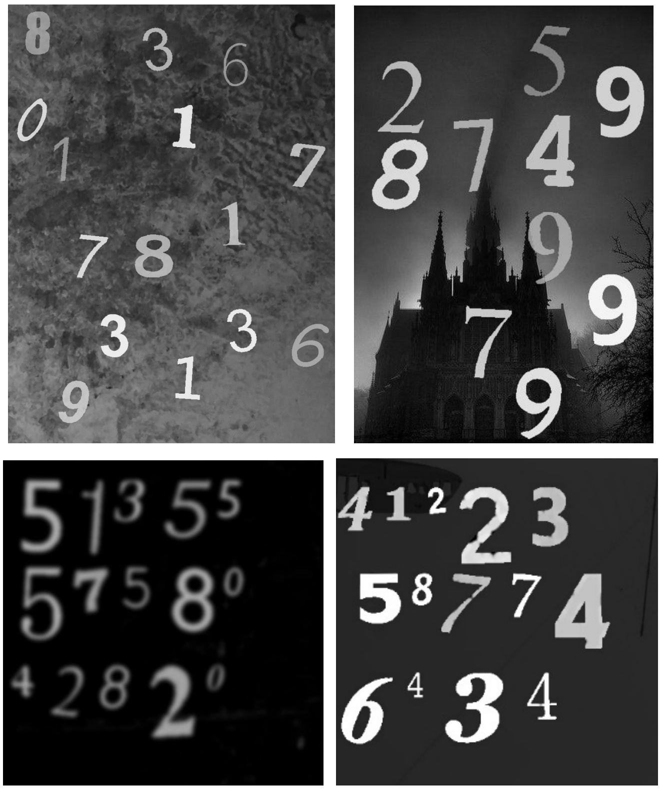

Fig. 3. Image samples from the generated data set

The training dataset is characterized by the following parameters:

- The color of the digits is changed to a random shade of gray in the range from 200 to 255;
- The digits are tilted at a random angle up to 40 degrees;
- After inserting the digits on the background, the final image is darkened by a random value from 0% to 50%;
- The final image is blurred by a random value from 0 to 50%;
- Digits are inserted into the background image and saved in folders with training and test images in a 4:1 ratio.

The training was performed in a Python environment based on the Keras and Tensorflow libraries on a PC running Windows 7 with the following characteristics: Intel Core i5-3450 / 16Gb RAM / 1000Gb HDD. The data set was divided into a training set (train set) of 3200 images and a test set of 800 images. All the main parameters of the neural network model were saved, as in the original model.

A stochastic gradient descent was chosen as the optimization method, with parameters learning rate = 5e-3, decay = 0.0005, momentum = 0.9. The loss function was taken the same as it was used in the original YOLO model. Transfer learning was applied: the initial weights were taken from a similar neural network trained on the COCO dataset.

The training was carried out for 100 epochs. In case 10 epochs presented no decrease in the value of the loss function on the test part of the dataset (Test set), then the training was interrupted. The graph of changes in the loss function in the learning process for 100 epochs is shown in Fig. 4.

The neural network was tested on 87 real counter images. Image data was recognized by a trained neural network, as a result of recognition all the digits were printed on the images with a recognition probability of more than 50%. The obtained examples of counter recognition are shown in Fig. 5.

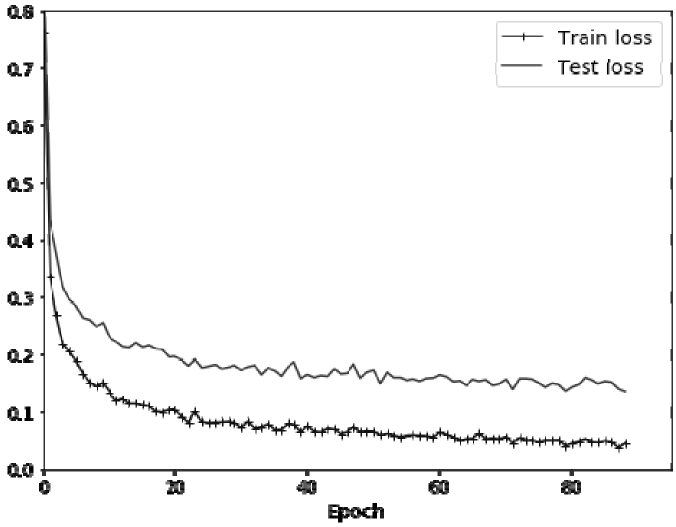

Fig. 4. Loss functions for training and testing datasets

The following indicators were taken as a measure of accuracy and completeness:

$$Precision = \frac{N_{all\ found} - N_{wrong}}{N_{all\ found}}, \quad Recall = \frac{N_{all\ found} - N_{wrong}}{N_{all} - N_{wrong}},$$

where $N_{all\ found}$ – all found digits of readings of counters; N_{wrong} – incorrectly recognized digits of meter readings; N_{all} – all digits of counters.

The results of the accuracy assessment are presented in Table I.

TABLE I. METERS READING RECOGNITION RESULTS

	Float values	Integers
N_{all}	520	434
$N_{all\ found}$	359	312
N_{wrong}	72	44
Precision	0.8	0.86
Recall	0.64	0.69

Fig. 5. Examples of real meters reading recognition

V. DISCUSSION

Based on the experimental results, the following conclusions were made on how to improve the training dataset:

1) It is necessary to differentiate the training dataset into three equal parts:

- without transformation;
- with slight dimming and blur;
- with strong dimming and blurring;

2) Blurring and dimming should be limited. A large number of darkened and blurry images in the dataset lead to poor recognition of readings under normal conditions.

3) An increase in the number of images in the dataset leads to an improvement in the performance of the neural network (see indicators Precision and Recall).

As a result of Generator changes, according to the conclusions made, the following datasets were received:

- 4K images, no dimming, blur all images;
- 8K images with a blackout and blur of all images;
- 10K images with a half of images with dimming and blurring;
- 20K images, one third is without dimming and blurring, one third with weak dimming and weak blurring, and one third with strong dimming and strong blurring.

As a result of training the neural network on the generated datasets, the accuracy rose to 96% and the recall was improved up to 90% (see Fig. 6, 7).

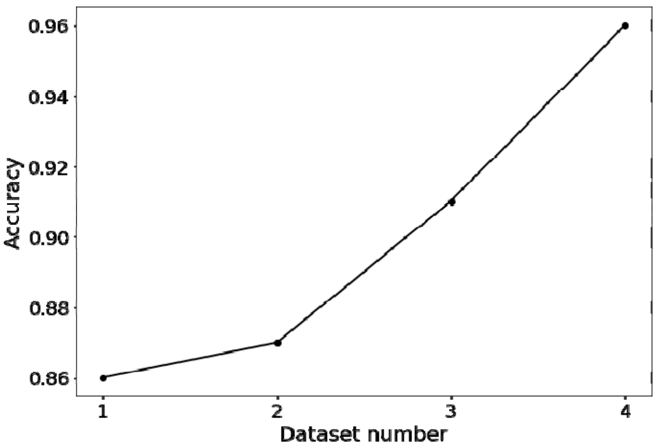

Fig. 6. Neural network accuracy

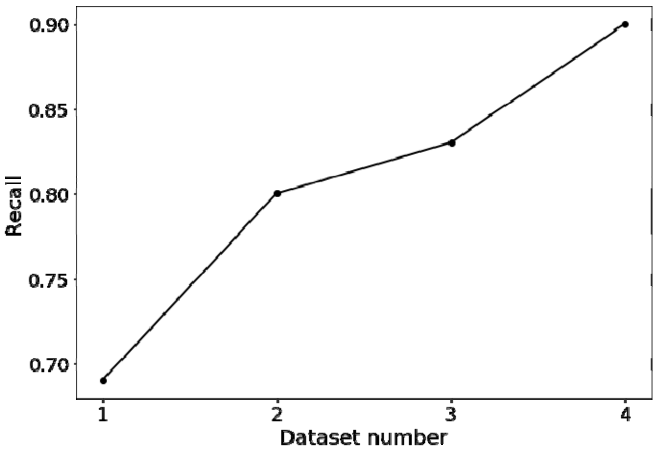

Fig. 7. Neural network recall

The new solution has the following features:

1) The readings are still not well recognized in low light conditions and / or with poor quality shooting;

2) Due to the different fonts on the counters themselves, the neural network can sometimes confuse similar digits (3 and 5, 1 and 7);

3) Implementation of the system on medium-power mobile devices running Android OS shows that the processing speed of one counter increases up to 1 – 3 seconds, which is unacceptable for real applications;

4) Although the achieved accuracy and completeness are acceptable, but the size of the scales is too large to be used on a mobile device (400 MB), it is necessary to reduce its size.

To solve the problems with speed and size in the convolutional layers of the neural network, the number of convolution kernels was reduced by 4 times, which led to a decrease in the parameters by about 16 times: from 50 million to 3.2 million. Decreasing the parameters made it impossible to use transfer learning. The default initializer in Tensorflow 1.13.1 (glorot_uniform_initializer) was used to initialize the scales. Other parameters of the neural network remained the same.

To improve accuracy and completeness, the datasets of 50K and 100K images were generated with the following process:

- The color of the digits is randomly changed to one of the shades of gray from 200 to 255;
- The digits are accidentally cut from the bottom or top by a value from 0% to 20%;
- The digits are tilted at a random angle from -30 to 30 degrees;
- After inserting the changed digits into the background, the final image is blurred by a random amount from 0% to 70%;
- After darkening, the final image is darkened by a random amount from 0% to 80%;
- Images are saved in folders with training and test images in a ratio of 80:20.

The results of training and recognition are presented in Table II and Fig. 8.

TABLE II. RESULTS OF RECOGNITION BY A NEURAL NETWORK BASED ON NEAR-ZERO SCALES OF BRIGHT DIGITS IN IMAGES WITH COUNTERS

	50 000 Images	**100 000 Images**
N_{all}	434	434
$N_{all\ found}$	302	357
N_{wrong}	57	72
Precision	0.81	0.80
Recall	0.65	0.79

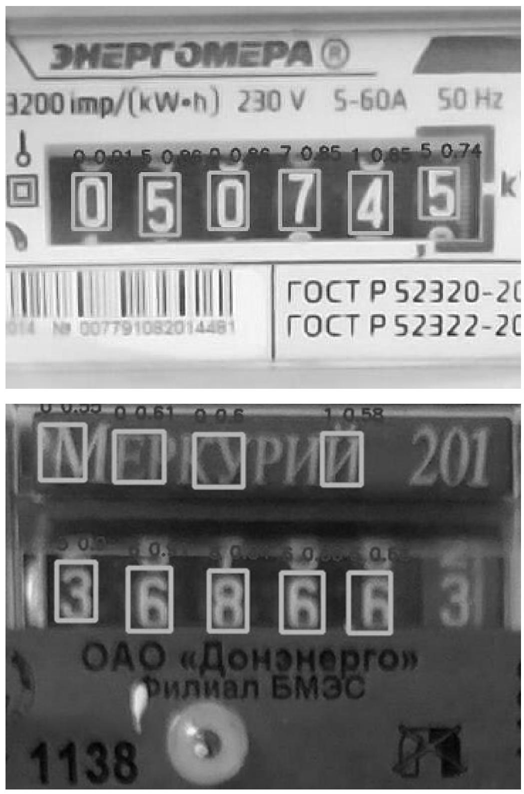

Fig. 8. Successful recognition results

According to the experimental results, an increase in the dataset to 100 000 images gives a significant increase in completeness, but the accuracy indicator is slightly decreased. Despite the lower accuracy comparing to the initial neural network, this solution can significantly increase the recognition speed, and lower the size of the scales from 400 to 24.5 MB, which allows deployment of a neural network on a mobile device.

The main recognition problems arise with photos of very poor quality (see Fig. 9). There can be many reasons for the poor image quality of the counter readings, for example, lack of focus when shooting, random camera movements at the time of shooting, glare from the sun, etc. Recognition results are presented in Fig. 10.

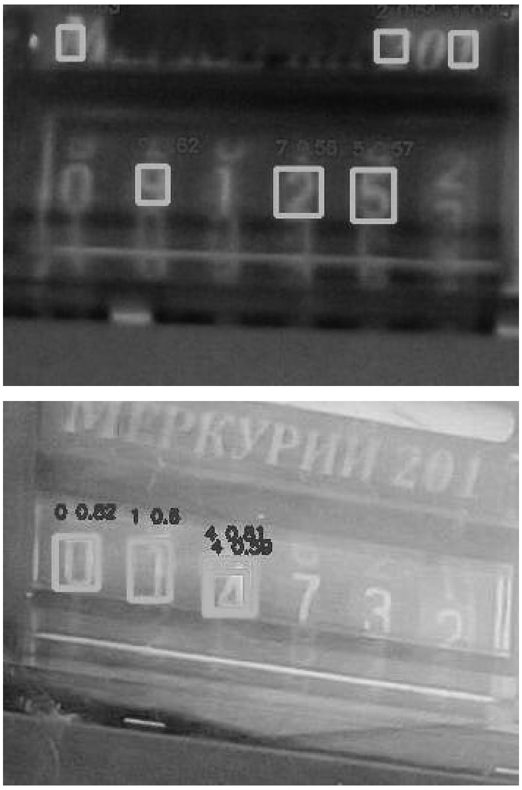

Fig. 9. Poor recognition example

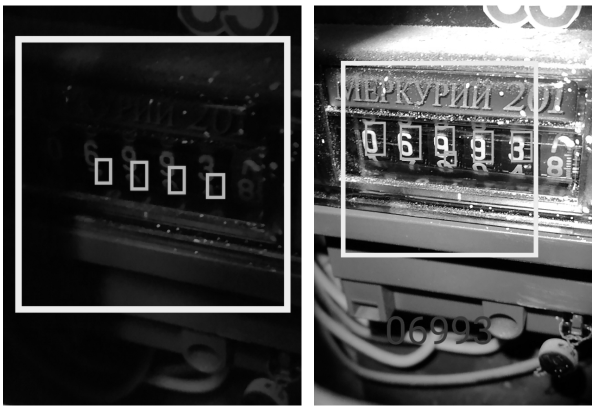

Fig. 10. Recognition started (left) and done (right) in 1 – 2 sec after start

Therefore, the Pre-launch AI matcher solution eliminates the drawbacks, allowing the user to move the mobile device to find the camera position when the readings on the counter are visible well enough for recognition

VI. Conclusion

Introduction of AI Matcher allows better targeting the neural network training process. After adding logic to extract readings, successful results were obtained in real applications using a fully synthetic dataset for training. To further improve the quality of recognition, it is planned to facilitate the model used, by reducing the number of layers and filters.

AI Matcher implementation for intelligent recognition has high prospects not only in the meters surveying but also in other problem domains that require flexible logic. The proposed approach allows building adaptive solutions with controllable sensitivity to the features of images considering the variety of problems being solved.

Acknowledgment

The paper was supported by RFBR, according to the research project № 20-08-00797.

References

[1] A. Ivaschenko, A. Krivosheev, and P. Sitnikov, "Multi-agent solution for a distributed intelligent photo surveying", *Proceedings of the 2019 European Simulation and Modeling Conference (ESM 2019)*, EUROSIS-ETI, 2019, pp. 73-78

[2] A. Ivaschenko and A. Krivosheev, "Distributed processing of electrical meters surveying", *Proceedigs of the 2020 Moscow Workshop on Electronic and Networking Technologies (MWENT)*, 2020, pp. 1-4

[3] D. Romero and F. Vernadat, "Enterprise information systems state of the art: past, present and future trends", *Computers in Industry*, Vol. 79, 10.1016/j.compind.2016.03.001, 2016

[4] N. Bessis and C. Dobre, "Big Data and Internet of Things: A roadmap for smart environments", *Studies in computational intelligence*, Springer, 2014, 450 p.

[5] M. Egmont-Petersen, D. de Ridder, H. Handels, "Image processing with neural networks – a review", *Pattern Recognition*, 35 (10), 2002, pp. 2279-2301

[6] B. Pt and P. Subashini, "Optimization of image processing techniques using neural networks – A review", *WSEAS Transactions on Information Science and Applications*, 8 (8), 2011, pp. 300-328

[7] M. Jena and S. Mishra, "Review of neural network techniques in the verge of image processing", *Advances in Intelligent Systems and Computing*, vol 628. Springer Singapore, 2018, pp. 345-361

[8] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, "Deep learning, vol. 1", MIT press Cambridge, Vol. 925, 2016

[9] S.P. Orlov and R.V. Girin, "Intelligent technologies in the diagnostics using object's visual images". *Studies in Systems, Decision and Control*, vol. 259. Springer Nature Switzerland, 2020, pp. 301-312

[10] Z. Zhao, P. Zheng, S. Xu and X. Wu, "Object detection with deep learning: a review", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, 2019, pp. 3212-3232

[11] A. Arcos-Garcia, J. Alvarez-Garcia and L. Soria Morillo. "Evaluation of deep neural networks for traffic sign detection systems", *Neurocomputing*, 316. 10.1016/j.neucom.2018.08.009, 2018

[12] Y. Pi, N. Nath and A. Behzadan, "Convolutional neural networks for object detection in aerial imagery for disaster response and recovery". *Advanced Engineering Informatics*. 43. 101009. 10.1016/j.aei.2019.101009, 2020