

# Prelaunch Matching Architecture for Distributed Intelligent Image Recognition

Anton Ivaschenko<sup>1</sup>, Arkadiy Krivosheev<sup>2</sup>  
Computer Technology Department  
Samara State Technical University  
Samara, Russia

Pavel Sitnikov<sup>3</sup>  
SEC "Open Code"  
Samara, Russia

**Abstract**—The paper presents a multi-agent solution for dynamic combination of several artificial neural networks used for image recognition. As opposed to the existing methods there is introduced a dispatcher agent that provides prelaunch matching of possible pro-active identification algorithms through competition. The proposed solution was implemented to solve a problem of stream processing of photo images produced by a number of distributed cameras using an intelligent mobile application. It was probated and utilized in practice to capture the results of electrical meters that are manually monitored by a group of patrol personnel inspectors using hand held devices. Prelaunch matching architecture allowed increasing the quality of digits recognition using various neural networks depending on the operating conditions.

**Keywords**—Multi-agent technology; artificial neural networks; image recognition; electricity meter data processing

## I. INTRODUCTION

Artificial neural networks are widely used for image recognition nowadays. However their application in practice is still concerned with the low versatility caused by filtering property that leads to the lack of multitasking. The image recognizing system needs to process a large variety of data in different operating conditions, and the better processing of one type leads to the failures in other cases. In addition to this the neural network is often being especially trained to reduce the noise affected by surroundings, which can be useful to identify the context sometimes.

Combination of several neural networks within the solid solution [1, 2] is a distinguished approach to improve the quality of recognition making the system adaptive to changing conditions. In this case a multi-agent paradigm becomes suitable and efficient to build an intelligent system with a distributed architecture [3, 4]. Despite the exploitability and potential capacity of this solution the ways of combination remain different, which makes it hard to develop a fully configurable and adaptive system in practice.

For example, electrical or gas meters surveying is so far concerned with visual monitoring of meter readings of various counters, which remains a challenging area for image recognition. This job is usually performed by professional patrol personnel aimed to check the visual status of the system, take the reading and snap a photo for validation. In order to succeed most of them use smartphones or tablet computers that could be upgraded with intelligent software.

Variance in meter types and operating conditions makes it problematic to use single neural network.

To cover this gap there was developed a prelaunch matching software solution based on the concept of multi-agent architecture of distributed intelligence. In this case the software agents are introduced not to simulate or be deployed on hand held devices of monitoring inspectors. Each agent represents an autonomous recognizer specializing on processing of a certain type of meters. Additional Matcher agent is supplemented to scope out various objects and assigns them to the most corresponding recognition modules. It can be based by an extra neural network itself or implement preliminary defined rules or reasoning.

## II. MATERIALS AND METHODS

Technologies of intelligent image recognition using neural networks and positive examples of their practical applications are presented in [5, 6]. Image analysis and pattern recognition remain the challenging areas of classification and clustering using the modern technologies of artificial intelligence [7, 8].

Neural networks provide adequate and stable identification of real objects and items with a complex shape. Textual data (words and numbers) are identified with a sufficiently high accuracy even in case of fuzzy or washed-out picture. More specific solutions are presented in [9, 10].

The quality of data processing and analysis can be improved under the context of monitoring procedure and results of preceding identification. Several data sets are combined to analyze multiple layers of a system at once. This approach is widely used for medical data analysis and can be disseminated for a cyber-physical system [11], which can interlink all related data sets (e.g., images, text, measured values, scans) and offer visual analytics to support experts.

Distributed image recognition can be considered as a problem of the Internet of Things. Combination of the Internet of Things as a major data source and Big Data technologies is a powerful tool for information processing and analysis [12, 13] being successfully used at modern enterprises with distributed structure, i.e. electrical networks and oil pipelines.

Modern software architectures improve the performance of data processing in real time using an approach of parallel computing, multi-agent modeling and distributed decision-making support [14, 15]. This approach offers a way of designing adaptive systems with decentralization over

distributed and autonomous entities organized in hierarchical structures formed by intermediate stable forms. Its implementation in practice requires development of new methods and tools for supporting fundamental mechanisms of self-organization and evolution similar to living organisms (colonies of ants, swarms of bees, etc.).

Multi-agent technology can significantly enrich the adaptability of an intelligent system by making it possible to add new components and thus increase the number of options considered. Such an extensive development does not require the changes in an existing logic of already deployed components. The proposed approach is based on the experience of development of distributed image recognition solutions [16 – 18]. This paper describes one further step in this direction.

### III. SOLUTION ALTERNATIVES CONSIDERED

To develop a software solution for distributed intelligent image recognition taking the problem of electrical meters surveying as an example, the following options were analyzed.

Initially, a centralized architecture was created and put into operation, but according to the results of the work, it turned out that mobile devices did not always provide stable communication with the central server. There were also performance problems at times when, in addition to professionals, company management sent other employees to collect data. Such an increase in requests had a negative effect on recognition time. To solve these problems, the architecture of distributed recognition of readings was developed.

Centralized recognition implementation initially seems to become an obvious solution considering the requirements of neural network study and functioning. The logic of the module itself (see Fig. 1) is based on a fairly simple linear architecture. Images of meter readings made by inspectors using smartphones with the Android operating system are transmitted via the Internet to the central computing server, where they are processed in single-threaded mode (in the Main Thread stream) using the PreprocessImagePipeline method.

This method processes the input image in more than 1000 different ways in order to find outlines of readings on the image. Those processed images on which potentially suitable contours were found are transferred to the ProcessImagePipeline method.

The ProcessImagePipeline method takes a closer look at the resulting contours and eliminates the excess ones. Using the remaining contours, it cuts out the numbers from the processed images and passes them for recognition to the RecognizePipeline method. Inside the RecognizePipeline method, digital images are recognized using the neural network described below. For each of the previous stages of image processing options, a recognition result is obtained. Among the whole set of results, the best result is selected.

Based on the results of the operation of the centralized recognition module, it was decided to transfer the recognition process from the central server to specialized autonomous

recognizers. These modules can be deployed either on inspectors' smartphones, or on dedicated servers on the cloud. The architecture of the distributed reading recognition module appeared is shown in Fig. 2.

In the main thread, the Controller method receives images from the camera, sends them to the UniversalRecognizer and Tracker methods, and also passes the results to the Agregator method. The Agregator method saves and analyzes the results, according to the results of the analysis, returns a report to the Controller method, which indicates whether the final result is ready or if you want to continue collecting results.

In a separate Recognition thread, the UniversalRecognizer method works, which is the centralized recognition module in which the number of processing of the input image is reduced to 10 and the type of processing, is randomly selected.

Another Tracking thread is running the Tracker method. It takes an input image and returns the difference with the previous image, i.e. direction in which direction the camera has moved. This information is subsequently passed through the Controller along with the recognition results to the Agregator method.

There, it helps to compare the results between themselves, obtained at different intervals, as well as reset the old results with large changes in the position of the camera.

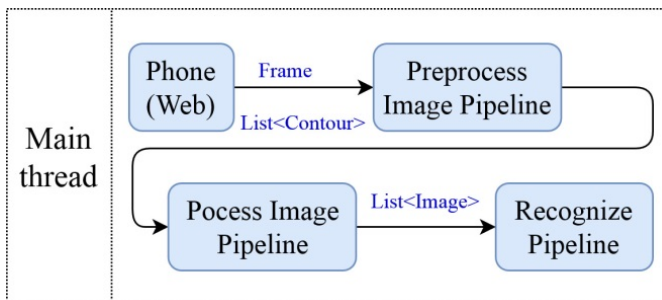


Fig. 1. Architecture of a Centralized Recognition Module.

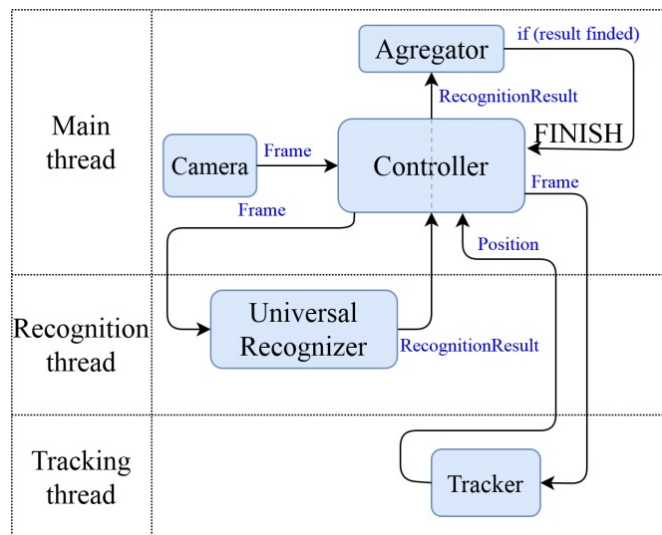


Fig. 2. Distributed Recognition Module Architecture.

#### IV. PRELAUNCH MATCHING ARCHITECTURE

The achieved results were generalized by a concept of prelaunch matcher. The main idea is to split the system to distributed parts with autonomous behavior. These parts proactively communicate looking for the best combination of services to solve the initial problem. For the perspective of neural networks implementation in practice this approach gives an opportunity to combine several solutions instead of training one neural network, which might require significant costs and 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. 3.

Considering the nature of the proposed approach it can be implemented using multi-agent technology as a software development paradigm. Data from camera is sent to the Matcher agent. Matcher agent operates on the basis of logic, neural networks and knowledge bases; its purpose is to choose the data computing strategy. If Matcher has selected several preprocessors, then each of them is questioned whether it is able to find his patterns on the sample from the received data, according to the results of the answers, the list of the preprocessors selected for processing is specified.

There are three types of Matcher agent strategies: auction, automatic dispatching, and competition:

- The auction is a survey of agents in order to find out which of them is capable of processing data. Based on the survey results, the Matcher agent selects Recognizer agents to process the current data set;
- When automatic dispatching, the Matcher agent independently chooses which agent to prefer for data processing;
- If the competition strategy is applied, then all available agents are involved in data processing, and the best ones are selected based on the processing results.

Agents use basic logic and neural networks in their work. Digit recognizers are used to find numbers in the image. Postprocessor tries to identify the digits that relate to the counter among all the digital symbols in view, the resulting data is transferred to the current storage. When the resulting data are found, the Frame tracker is connected, its purpose is to calculate the movements of the camera between the frames.

This data allows the Matcher agent to compare readings taken at different points in time. Also, as a result of moving the camera away from the counter, the Current storage may be reset. Over time, data from the counter taken from different angles accumulate in the current storage. Matcher selects the best counter results from all successful frames and generates the final recognition result.

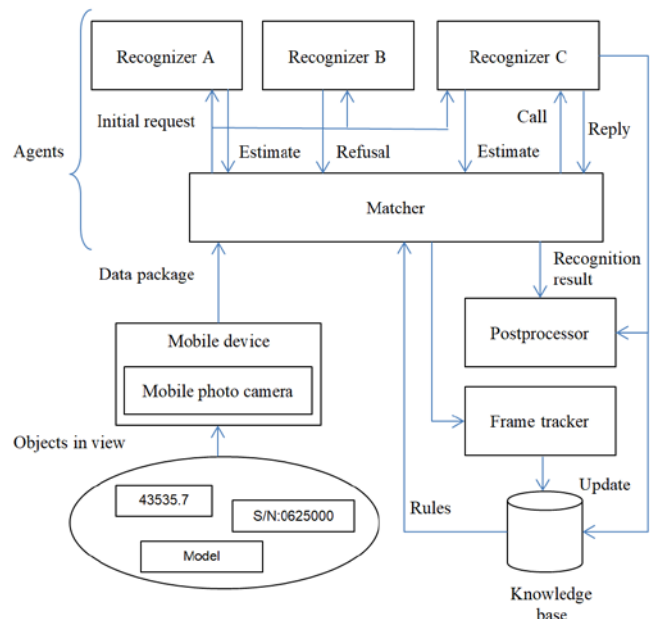


Fig. 3. Prelaunch Matching Architecture.

Auctioning strategy allows distributing the logic of decision-making between the components of the multi-agent architecture. When meeting an unknown object the system can organize a survey sending the requests to all the involved recognizing agents. They can either reply with and accepting of refusing answer based on preliminary understanding of the object type, or try to perform image recognition and reply in case of good quality of identification. After receiving the answers the Matcher can choose the best one and start negotiating with the corresponding recognizer on further identification.

Implementation of the multi-agent approach allows providing high autonomy of the recognizers, introduce new recognition algorithms with minor architecture changes and mix them in case of high level of uncertainty. Rather than other multi-agent implementations of distributed intelligent applications this solution does not require differentiation between the scopes of neural networks. Several different recognizers can be trained using the intersectional sets. Therefore such architecture remains open and provide an opportunity for permanent development by adding new recognizers without replacing the previous ones.

#### V. IMPLEMENTATION IN DISTRIBUTED PHOTO SURVEYING

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 both analog and digital. The vast majority of them cannot transmit the values electronically and require photo surveying.

To solve this problem there was developed special software for hand held devices, tablets and smartphones (see Fig. 4) 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.

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.

The described above approach was implemented, probed and tested for a convolutional neural network based on the LeNet architecture used to recognize the number symbols. When initializing the model weights, the Xavier Initialization was used.

For convolutional layers, the IDENTITY activation function was set, for a fully connected layer – RELU, for the Output layer – SOFTMAX. To exclude retraining, regularization L2 with parameter 0,0005 was used. The learning speed was set depending on the current iteration: every 200 steps until the 1000th iteration, the speed sequentially decreased from 0.06 to 0.001, after 1000 iterations the speed did not change.

An initial attempt was made to train a neural network based on a set of handwritten digits MNIST. Soon after the start of the work it became clear that it was not possible to get a good result with it. To create a dataset, about 1000 fonts were collected; based on these fonts, images of numbers in the amount of 10,000 copies were generated.

After augmenting this set with rotations and shifts, a dataset was created consisting of 196,000 images of digits, see Fig. 5.

It turned out that 30% of all digits “1” did not differ from each other, they were replaced by additional transformations of the remaining original “1”. The neural network was trained on the received dataset. The results of training the neural network are presented in Fig. 6.

To check the quality of the modules, a test kit was assembled, including 138 images of digital meters (777 digits in total) and 95 images of analog meters (534 digits in total). Because the distributed recognition module receives a video stream as an input, then a sequence of images was simulated for it by means of small offsets of the tested photo. Recognition results on a full set of images will be as follows (see Table I).

Some of the images in the test set were of poor quality, which is even difficult for a person to parse the readings. If you sort the set by image quality and choose the best half, the results can be improved.

One can see that implementation of a multi-agent architecture allows increasing the quality of digits recognition as a part of a distributed intelligent photo surveying solution.



Fig. 4. Application user Interface.

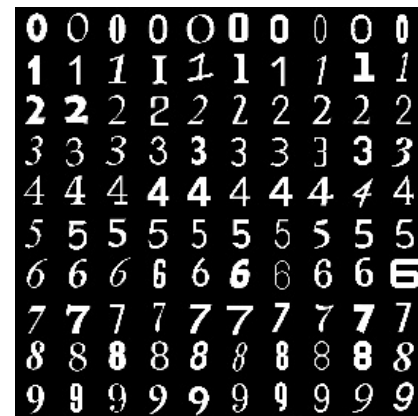


Fig. 5. Updated Training Dataset Images.

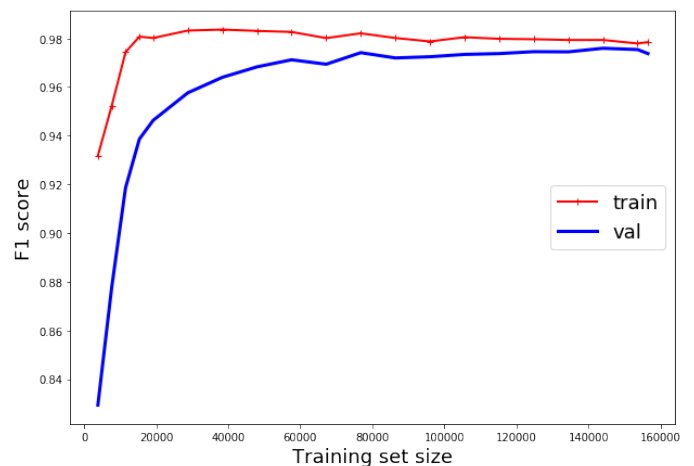


Fig. 6. Dependency of F1 from the Volume of the Dataset.

TABLE I. RECOGNITION QUALITY OF DIFFERENT SOLUTIONS

	Centralized recognition module (FULL dataset)	Distributed recognition module (FULL dataset)	Centralized recognition module (realistic dataset)	Distributed recognition module (realistic dataset)
Analog meters	392/534 = 73.4%	432/534 = 80.9%	197/212 = 92.9%	203/212 = 95.8%
LCD meters	611/777 = 78.6%	640/777 = 82.4%	368/378 = 97.3%	374/378 = 98.9%

## VI. APPLICATION TO PRACTICE

The proposed solution was used in the specialized mobile application for photographing the readings of electricity meters, their transmission to the data processing center, recognition and operational analysis by the staff of a regional energy distribution company.

Search for words was carried out by Tesseract. By the relative position of the words, there was a meter mask, by the absolute position of the words in the photo there was an area with readings according to the data obtained from the mask.

As an alternative, a video stream from the phone's camera was implemented in the browser, a frame was displayed on top of the video stream, the user needed to point the phone at the meters that the readings fell into the frame, and then click on the recognition button.

Recognition works according to the following algorithm. The inspector takes a photo; it is processed by a series of filters that increase the quality and sharpness. Then the color image is converted to b / w in about 50 times in different ways, each image is subjected to the following actions:

- contours are searched for in the image, contours are outlined by rectangles;
- among the obtained rectangles, sequences are found that lie on one straight line and having the same size, among all sequences, the best one is taken according to an empirically selected formula with arguments "rectangle size" and "number of rectangles";
- according to the rectangles of the b / w image, sections with potential readings are cut out;
- the cut out images are processed and recognized, the recognition results are saved.

The results of implementation allowed performing the series of experiments and probing the solution in practice.

## VII. CONCLUSION

Prelaunch matching architecture provides additional benefits from combining several neural networks into a solid intelligent solution for distributed photo surveying. As opposed to other solution it allows integrating several recognizers trained using the intersectional sets, which makes it open for permanent development by adding new recognizers without replacing the previous ones.

The results of implementation, testing and practical use illustrate the benefits of autonomous pre-processing of photo images using the mobile application with a multi-agent architecture. Next steps are considered with the study of possible agent strategies aiming an increasing the recognition quality in various conditions.

## ACKNOWLEDGEMENT

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

## REFERENCES

- [1] N. Ueda, "Optimal linear combination of neural networks for improving classification performance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 2; 2000, pp. 207-215
- [2] L. Lenc, P. Král, "Combination of neural networks for multi-label document classification", *International Conference on Applications of Natural Language to Information Systems*; 2017, pp. 278-282
- [3] A. Quteishat, C.P. Lim, J. Tweedale, L.C. Jain, "A neural network-based multi-agent classifier system", *Neurocomputing*, vol. 72, issues 7-9; 2009, pp. 1639-1647
- [4] R. Asadi, N. Mustapha, M. Sulaiman, "A framework for intelligent multi agent system based neural network classification model", *International Journal of Computer Science and Information Security*, vol. 5, no. 1; 2009, pp. 168-174
- [5] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, *Deep learning*, vol. 1, MIT press Cambridge; 2016
- [6] M. Egmont-Petersen, D. de Ridder, H. Handels, "Image processing with neural networks - a review", *Pattern Recognition* 2002; 35 (10), pp. 2279-2301
- [7] B. Pt, 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
- [8] M. Jena, 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
- [9] Z. Zhao, P. Zheng, S. Xu, 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
- [10] A. Arcos-Garcia, J. Alvarez-Garcia, L. Soria Morillo, "Evaluation of deep neural networks for traffic sign detection systems", *Neurocomputing*, 316. 10.1016/j.neucom.2018.08.009; 2018
- [11] A. Holzinger, "Interactive machine learning for health informatics: when do we need the human-in-the-loop?", *Brain Informatics 2016*, Volume 3, Issue 2. pp. 119-131
- [12] N. Bessis, C. Dobre, "Big Data and Internet of Things: A roadmap for smart environments", *Studies in computational intelligence*, Springer 2014. 450 p.
- [13] G. Fortuno, P. Trunfio, "Internet of Things based on smart objects: technology, middleware and applications", Springer-Verlag New York Inc., 2014, 250 p.
- [14] M. Wooldridge, *An introduction to multi-agent systems*, John Wiley and Sons, Chichester; 2002, 340 p.
- [15] V.I. Gorodetskii, "Self-organization and multiagent systems: I. Models of multiagent self-organization", *Journal of Computer and Systems Sciences International* 2012, vol. 51, issue 2, pp. 256-281
- [16] A. Ivaschenko, A. Krivosheev, "Distributed processing of electrical meters surveying". 2020 Moscow Workshop on Electronic and Networking Technologies proceedings; 2020. pp. 1-4
- [17] A. Ivaschenko, A. Krivosheev, P. Sitnikov, "Multi-agent solution for a distributed intelligent photo surveying". Proceedings of the 2019 European Simulation and Modeling Conference 2019. pp. 73-78
- [18] Ivaschenko A., Sitnikov P., Surnin O. *Accented visualization for Augmented Reality // Emerging Topics and Questions in Infocommunication Technologies*, Cambridge Scholars Publishing, 2020. - pp. 74 - 97