Neural network model of machine parts classification by optical scanning results

V A Pechenin, M A Bolotov and E Yu Pechenina
Samara University, Institute of Engines and Power Plants, Moskovskoe shosse, 34, Samara, Samara Region, 443086, Russia

E-mail: vadim.pechenin2015@gmail.com

Abstract. The article considers the solution to an automation problem for assembly processes associated with determining a type of the part delivered for assembling. The parts delivered for assembling are preliminarily measured with an optical scanner. In order to solve the problem for determining the part type, convolution neural network organization was matched with the data classification. Stl-patterns for three turbine rotor parts were taken, the training and test samples were simulated, several convolution neural networks were trained and the optimal parameters ensuring 100% classification accuracy were selected.

1. Introduction

The important direction in the robotic systems development is to increase their activity independence. A significant number of robotic systems for industrial use, involved in the technological stages of welding, painting, completing and assembling, etc., are characterized by clearly defined operation algorithms, which have limited flexibility and adaptability to changing production conditions. Ensuring full activity independence for robotic systems to implement the assembly cycle requires the development of intelligent control systems, which shall complete two tasks at least: 1) to define type of the part delivered for assembly and surface components for their spatial orientation; 2) to predict geometrical parameters forming while assembling details and defining quality of the product to be assembled [1, 2].

The part surface connection process (pairing) complexity creates necessity for wide use of numerical methods. When solving assembly tasks, the geometry shall be measured using coordinate control means [3]. In this case, the coordinate measuring instruments accuracy becomes very important [4, 5, 6] for a successful assembly. In solving the task to identify the parts automatically and to recognize the facets involved in the digital process, it is possible to use different algorithms to work with the geometry [7, 8]. To automate the search for initial conditions (data with which the algorithms begin to work while recognition), a large number of approaches are developed now to recognize three-dimensional objects based on computer vision [9, 10]. The most approaches are based on the convolutional neural network use [11].

The work aims to solve the task for classifying the parts to a known type after their measurement using optical scanners.
2. Neural network model for data classification after scanning

In order to classify any objects, a matrices set containing information about the picture pixel values is delivered into the input to the convolutional neural networks. In the case of computer vision, one or more cameras receive a detail images set. The optical scanning results in a 3D facet model in *.stl format. In order to build the data used to classify the part, three projections of the stl-pattern on the coordinate planes are created: XOY, YOZ and XOZ. Thus, three images in Grayscale, i.e. three number matrices that are input “layers”, are delivered into the network. The organization of the convolutional neural network for classifying the measured parts is shown in figure 1.

![Figure 1](image)

**Figure 1.** The convolutional neural network organization for parts classification.

The network is a connected layers sequence starting with a convolutional layer and ending with the SoftMax layer. The data are normalized in the range from zero to one. Let’s give a brief description for the concerned network layers.

The convolutional layer includes a filter for each channel, whose convolutional core processes the previous layer by fragments, summarizing the matrix product results for each fragment. The convolution kernel weight ratios, represented by a small matrix, are unknown and are set during training.

The scalar result of each convolution falls on an activation function, which is a non-linear function. The activation layer is usually logically combined with the convolutional layer. The piecewise linear function ReLu [11] is selected. The process essence for batch normalization is described, for example, in [12].

Pulling layer (otherwise sub-sampling, down-sampling) is a non-linear condensation of the feature map, with a group of pixels (usually 2×2 in size) compacted to one pixel through a non-linear transformation. The most common is the maximum function, which is used in this network.

Dropout is a method for regulating neural networks [13, 14]. Each neuron, except the most recent output layer, is set to a certain probability \( p \), which it will be thrown out of the network with. After several image convolutions and compaction with pooling, the system is rebuilt from a specific high-resolution pixel grid to more abstract feature maps, and eventually, a large set of channels remains that store a small data amount (even one parameter), which are interpreted as the most abstract concepts identified from the original image. These data are combined and transmitted to a normal fully connected neural network which can also consist of several layers. At the same time, fully connected layers already lose the spatial structure of pixels and have a relatively small size (relative to the number of pixels of the original image). SoftMax layer (soft maximum function). This layer normalizes the previous layer results in such a way that at its output the probabilities of object’s
relation to the considered classes will be formed [15]. That is, at the output from the softmax layer, \( n \) of numbers equal to the number of classes defined while network creating (in our case, three) appears, varying from zero to one. The ordinal number of the largest number determines the class which the object belongs to.

The share indicator of correctly classified objects \( N_{\text{corr},c} \) in the total number of sampling objects may be used to evaluate errors in the classification task \( N_{\text{total}} \):

\[
\delta_{\text{class}} = \frac{N_{\text{corr},c}}{N_{\text{total}}},
\]

Accordingly, the network efficiency is 100\% if the coefficient is equal to 1.

3. Results
To train neural networks, parts sampling was simulated by simulating surface points and forming faceted bodies. A training sample consisting of 4000 objects and a test sample consisting of 400 objects was created. In the sample there were parts of the shaft, retainer and disk, stl-patterns of which are shown in figure 2.

For each of the objects, three projections-images in Grayscale with 112×112 pixels were made, an example for one of the "Disk" parts is shown in figure 3. Samples of stl-objects and projections were sampled in an algorithm implemented in the MATLAB environment. The neural network for classification was implemented in the Python language.

![Figure 2. "Disc" (a), "retainer" (b) and "shaft" (c) stl-patterns.](image-url)
Figure 3. "Disc" stl-pattern and its three projections.

A hundred percent efficiency ($\delta_{\text{class}}=1$) was shown by a network with three convolutional layers, at 50 training periods. Table 1 shows the convolutional neural network organization with numerical parameters showing 100% efficiency.

Table 1. Parameters of convolutional neural network layers for classification

| Layer type                | Output layer       | Pitch |
|---------------------------|--------------------|-------|
| Convolutional layer       | (None, 4, 112, 112)| 112   |
| Batch Normalisation       | (None, 4, 112, 112)| 448   |
| Activation of ReLU        | (None, 4, 112, 112)| 0     |
| MaxPooling layer          | (None, 4, 56, 56)  | 0     |
| Convolutional layer       | (None, 8, 56, 56)  | 296   |
| Batch Normalisation       | (None, 8, 56, 56)  | 224   |
| Activation of ReLU        | (None, 8, 56, 56)  | 0     |
| MaxPooling layer          | (None, 8, 28, 28)  | 0     |
| Convolutional layer       | (None, 16, 28, 28) | 1168  |
| Batch Normalisation       | (None, 16, 28, 28) | 112   |
| Activation of ReLU        | (None, 16, 28, 28) | 0     |
| MaxPooling layer          | (None, 16, 14, 14)| 0     |
| Convolutional layer       | (None, 32, 14, 14)| 0     |
| Batch Normalisation       | (None, 32, 14, 14)| 56    |
| Activation of ReLU        | (None, 32, 14, 14)| 0     |
| Dropout Layer             | (None, 32, 14, 14)| 0     |
| Fully connected layer     | (None, 6272)       | 0     |
| Softmax layer             | (None, 3)          | 18819 |
| Output values             | (None, 3)          | 0     |
The convolutional layer parameter: the step defines a step for the convolution sliding window. Thus, the network contains four convolutional layers, three pulling layers.

4. Conclusions
The presented work is devoted to development of the fast parts classification mechanism for the purpose of robotic systems independence while assembly cycle. In order to solve the classification task, a neural network model was used, an algorithm for preparing data to enter the neural network was developed. The received results lead to the conclusion that using of the matched neural network model allows to classify three parts with sufficient precision. The next research stage must be the recognition problem solution of the necessary edges for the certain type parts assembly. In addition, the developed approach application for more assembly parts will be considered later.

Acknowledgement
The work has been supported financially by the Ministry of Science and Higher Education of the Russian Federation within the framework of the Presidential Scholarship (CII-262.2019.5).

References
[1] Bolotov M A, Pechenin V A and Murzin S P 2016 Method for uncertainty evaluation of the spatial mating of high-precision optical and mechanical parts Computer optics 40(3) 360-9
[2] Bolotov M A, Pechenin V A, Ruzanov N V and Balyakin D M 2019 Neural network model in digital prediction of geometric parameters for relative position of the aircraft engine parts CEUR Workshop Proceedings 2416 87-94
[3] Zakharov O V, Bobrovskij I N and Kochetkov A V 2016 Analysis of Methods for Estimation of Machine Workpiece Roundness Procedia Engineering 150 963-8
[4] Yalovoy O A, Zakharov O V and Kochetkov A V 2015 The Centerless Measurement of Roundness with Optimal Adjustment IOP Conf. Series: Materials Science and Engineering 93 012024
[5] Rezchikov A F, Kochetkov A V and Zakharov O V 2017 Mathematical models for estimating the degree of influence of major factors on performance and accuracy of coordinate measuring machines MATEC Web Conf. 129 01054
[6] Zakharov O V, Balaev A F and Kochetkov A V 2017 Modeling Optimal Path of Touch Sensor of Coordinate Measuring Machine Based on Traveling Salesman Problem Solution Procedia Engineering 206 1458-63
[7] Roberts L G 1965 Machine perception of three-dimensional solids In J. Tippett et al., editors, Optical and Electro-Optical Information Processing, MIT Press, Cambridge 159–97
[8] Stepanenko I S, Pechenin V A, Ruzanov N V and Khaimovich A I 2018 Technique of increasing the accuracy of GTE parts manufactured by selective laser melting Journal of Physics: Conference Series 1096(1) 012143
[9] Hubei D H 1988 Eye, Brain, and Vision (New York: Scientific American)
[10] Kolb H E et al. 2013 Webvision: The Organization of the Retina and Visual System Retrieved from: http://webvision.med.utah.edu/book/
[11] Nikolenko S I, Kadurin A and Arkhangelskaya E 2018 Deep Learning (St. Petersburg: Peter)
[12] Ioffe S Y and Szegedy C 2015 Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift Proc. 32nd ICML pp 448-56
[13] Srivastavaetal N 2014 Dropout: A Simple Way to Prevent Neural Networks from Overfitting Journal of Machine Learning Research 15 (1) 1929-58
[14] Hinton G E Srivastava N, Krizhevsky A, Sutskever I and Salakhutdinov R R 2012 Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors arXiv 1207.0580v1 – 1–18
[15] Nikitin M U, Konushin V S, Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment Computer Optics 42 (5) 732-2