Selection of appropriate architecture and parameters of neural network for images recognition and classification

V V Kukartsev\textsuperscript{1,2}, A S Mikhail\textsuperscript{1}, A V Tarasevich\textsuperscript{1}, V S Tynchenko\textsuperscript{1,2}, A R Ogol\textsuperscript{2} and V V Khramkov\textsuperscript{1}

\textsuperscript{1} Siberian Federal University, 79, Svobodny pr., Krasnoyarsk, 660041, Russian Federation
\textsuperscript{2} Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation

E-mail: vlad_saa_2000@mail.ru

Abstract. In this article the image recognition and classification problem is considered. Further, there will be suggestion of the new way of using a convolutional neural network to solve the considering problem in an example of a particular task, such as the recognition of ischemic stroke on magnetic resonance images. Moreover, it is considering reasons of choosing the image classification and making arguments obtained by a literary analysis of studies affecting the task. Authors have investigated different neural networks architectures and the accuracy depending on training speed, numbers of layers, numbers of epochs and mini sample size. Experiments result is presented in the table form. During the studying, magnetic resonance images of different diseases having same signals with a considering pathology in diffusion-weighted images format has been selected for the training of the convolutional neural network and for getting the most favourable result. In the result, a suitable architecture with defined parameters has been selected to get the highest accuracy.

1. Introduction

There are a lot of image classification methods. However, the convolutional neural network is the best image processing algorithms in accuracy and speed as well as its modification now [1], [2], [3]. The choosing current parameters of neural network and its architecture were due to required object features to recognize [4], [5], [6]. Often the recognition process is complicated since investigated object properties are set without mathematical parameters. As an experiment, there is a choosing a task to recognize ischemic stroke on magnetic resonance imaging (MRI) pictures which are provided by MRI centers in Krasnoyarsk.

Nowadays the problem is relevant because the pathology as insult is the one of leading causes of death from circulatory system diseases. According to the World Health Organization, the insult ranks the third place of people’s death in the world. Under the Northwestern State Medical University named I.I. Mechnikov, in the Russian Federation there are 450000 people suffer from insult per year, 80% of these insults are ischemic stroke [7].

During the example, in the studying the developed neural network will speed up the process of detecting the ischemic stroke in several times and, also increase the accuracy of the diagnosis by a training neural network with various data sample.
2. Problem statement
The main studying object is the choice of the most suitable architecture in which there will be found an optimal balance of computational complexity and desired accuracy and, also selecting parameters for which the received neural network will produce the greatest accuracy.

3. Theoretical research
Image processing is any form of information processing with an image as input data. The image processing might be for getting image in the input as well as getting other information, for example, the text recognition, the image classification, etc. [8]

Image processing is in several steps. On figure 1, there is stepwise disassembling of the processing algorithm. The image classification operation usually precedes image processing for making conditions to increase the efficiency and the quality of image processing.

![Picture processing algorithm](image)

Figure 1. Picture processing algorithm.

On the next step, it is necessary to choose an algorithm of recognition and classification. The neural network is the distributed system of artificial neurons to process information [9].

The one of main advantages of neural networks is its learning ability. The appearing of new data in traditional algorithms induces its recount necessitate, it is not usual for neural networks, due to the training system, it adapts to different changes, using new data to discover a solution of the task each time.

In the way, neural network is pretty powerful and flexible instrument to solve different tasks in many scientific areas. However, the using classic neural network to recognise images is complicated usually because of the large vector dimension of input values in neural network, the large amount of neurons in intermediate layers and the high cost of computing resources for training and computing network [10].

For its part, in convolutional neural networks a key concept is “shared” weights, that is, a part of neurons in the considering layer of the neural network might be used by same weighting coefficients. Neurons, which use same weights, are combined into feature maps, and each neuron of it is associated
with a part of neurons in the previous layer. When network is computing, each neuron does the convolution operation in some area of the previous layer. Besides, except convolutional layers in the convolutional neural network there may be downsampling layers (it is layers which perform functions of reducing the dimension of the feature map space) and fully connected layers. All three types of layers can be alternated in random order, it allows making feature maps from feature maps, but in the practice, it means the ability to recognize complicated feature hierarchies.

In this way, advantages of convolutional neural network are [11]:

- The reducing the amount of training parameters and the increasing of the training speed compared to fully connected neural network.
- The resistance of the object shifting in input data. During the training, a convolutional neural network is shifting by parts of object. Therefore, training features do not depend on a position of “important parts”. The property of convolutional neural network helps to rise the classification quality [12]. It is the relevant problem because the objects on images can be in any part of the image.

4. Experimental research

The developing of the algorithm was performed in Python with JupyterNotebook, JetBrains PyCharm Community Edition and ML-framework TensorFlow.

An initial data for neural network training was MRI images in DWI format (diffusion-weighted images) because according to researches [13], on this type of images in a sharp phase of ischemic stroke 84.6% of patients have hyper intensive MR signal in lesion area. Otherwise, the using diffusion-weighted images can early and accurately determines areas of an acute cerebral ischemia.

For the training neural network, there were token images of several diseases, which have got same features on DWI:

- Glial tumors.
- Encephalitis.
- Pontine myelinolysis.
- Abscess.
- Hemorrhagic stroke.
- Glioblastoma.

On the figure 2 there is the demonstrating how similar there diseases are in diffusion-weighted images.

![Figure 2. Pictures of diseases: 1 - glioblastoma, 2 - encephalitis, 3 - ischemic-stroke.](image)

To sum up, for the training of neural network there are 2664 images, 112 images of it are for checking and 112 are for testing.
5. Image pre-processing

Image pre-processing has been performed in several steps. Whereas images have been provided from different resources, it has to be likely to the one size. To reach it, at the first step, images were scaled to 160x160 pixels with PIL (Python Imaging Library). Also, for images there was using of the interpolation method of Lanczos filter [14]. The way gives a great smooth result but time of processing increases.

The idea of filter is based on the applying of normalized function with the extended along axis \( x \) with main petal and the equal to zero outside the preassigned width parameter of the \( A \) interval.

The method is associated with the Lanczos window function: \( L_\omega(x) \), which is the main petal of function, outside the petal the window function is zero:

\[
L_\omega(x) = \text{sinc}(x/a).
\]  

While images are processing, since it is two-dimensional functions, for the convolution it is used two-dimensional Lanczos core:

\[
L(x,y) = L(x)L(y).
\]

The using this filter allows to get the high image sharpness.

On the second step of preprocessing, the noise removal is necessary because random artifacts and noise can aggravate readability and image quality. Whereas the suppression of white and Gaussian noise is more affective when linear filters are used, for system there is token medium filter [15] from OpenCV library with cv2.medianBlur() OpenCV function. It replaces each pixel to average value of surroundings in input image. In the result, as input image there is smoother image, which has got more clear borders of brain pathologies.

6. Choosing the neural network architecture

Factors affecting on the ischemic stroke recognition in MRI images are complicated. Therefore, the making an adequate mathematic model of the pathology recognition process allowing to diagnose the ischemic stroke with enough accuracy is difficult.

To solve the task there was performing a number of testing some neural networks with ready-made weights and different parameters on a processed data set. The test result is in table 1.

| Parameters | Using predefined layers | ResNet50 | GoogleNet | VGG16 |
|------------|-------------------------|----------|-----------|-------|
| Learning rate | 1e-5 | 1e-5 | 1e-5 | 1e-5 |
| Batch size | 20 | 20 | 20 | 20 |
| Epochs | 30 | 5 | 7 | 7 |
| Accuracy | 61% | 56.5% | 73.2% | 79.29% |

At the same time, the neural network has converged just in 50% of cases. According results, there was concluding that the using of VGG16 gives the best results in bigger data sample.

VGG16 [16] is the one of the most famous models. The architecture of VGG16 is in figure 3.
In input of the first convolutional layer, conv1 RGB images in 224x224 size are supplied. Onwards, images go through a stack of convolutional layers and three fully connected layers. The last layer is a soft-max layer.

All hidden layers have got “rectifier” non-linear function (ReLU) [17]. The function of activation allows reducing training time of neural network [18].

On the next step, there is the necessary to find out a suitable optimizer to solve the task. Nowadays, to solve the optimization parameter problem of neural network the large amount of different methods is developed [19]. Basically, these methods are first order methods and particularly variations of gradient descent. Weights of neural network are updated by applying iterative method of the back-propagation error. There was considering popular methods as a gradient descent and an Adam method.

The main idea of the gradient descent method is in the going to the direction of the fastest descent but it is set by the anti-gradient.

The Adam method combines the idea of activity accumulation and the idea of weaker weight update for typical features.

The most effective optimization method is the Adam method because it has the highest speed of convergence compared other methods. The simple gradient descent has enough small coefficient of a training step; therefore, the method is effective just at initial iterations and often gets stuck at local minima.

On the next step, it needs to find out a size of mini-sample. The size of it defines the amount of elements, which will be spread in network. During the study, the neural network has trained in different size of mini-sample. With increasing of the amount of data in mini-sample the accuracy is reducing, because the common size of sample is an enough large. Therefore, the conclusion is it is necessary to choose mini-sample of 20 elements.

Also, it needs to select enough amounts of eras. Studying results are presented in table 2.

Then result has been analyzed, there is the conclusion that on the sixth era the neural network begins to retrain because there was not enough data to continue successful training.

On figure 4 there are graphs of the accuracy dependence (the left one) and error (the right one) on the amount of iterations.
Table 2. Results of neural network testing with different number of epochs.

| Epochs | Convergence | Accuracy |
|--------|-------------|----------|
| 4      | yes         | 76%      |
| 5      | yes         | 77.29%   |
| 7      | yes         | 79%      |
| 8      | no          | 75%      |

Figure 4. Dependency graphs of accuracy and error against the number of iterations.

The accuracy, which the neural network gives, is 79%. However, the difference from the error percentage and accuracy of training and testing data set is still quite large, it is 20%. Therefore, VGG16 is applying for the image processing in 224x224 pixels, 160x160 pixels, in the last sixteenth layer there is the convolution to 2x2 size. According the conditions, the recognition process is sophisticated. Because of it, it was decided to take an architecture based on VGG16 architecture, however, it needs to set less amount of layers. In the VGG16 there are 13 convolutional layers. During the research, an amount of layers were gradually reduced and tested in different amount of eras to achieve result that satisfies tasks: the difference between an error and accuracy in training and testing samples have to be minimal but accuracy has to be maximal in the conditions. In table 3 there is the study result.

Table 3. Results of neural network testing with different number of epochs and layers (accuracy %).

| Epochs | Layers  | 4     | 5     | 6     | 7     | 8     |
|--------|---------|-------|-------|-------|-------|-------|
| 4      |         | 64    | 71.2  | 75.16 | 83.04 | 85    |
| 5      |         | 67.4  | 72.5  | 81.56 | 83.84 | 84.2  |
| 6      |         | 71.14 | 79    | 84.82 | 82.14 | 83.35 |
| 7      |         | 73.56 | 77.96 | 81.15 | 78.11 | 82.86 |
| 8      |         | 74    | 74.13 | 83.82 | 81.6  | 78.84 |
| 9      |         | 78.6  | 71.8  | 75.38 | 80.9  | 79    |
| 10     |         | 71.2  | 76.15 | 82.06 | 84    | 84.15 |
| 11     |         | 79    | 82.35 | 79.19 | 80.35 | 86.4  |

In this way, the best result has been get on 11 layers and 8 eras, 4 layers and 8 eras and also on 6 layers and 6 eras. However, according the difference between an error and accuracy in the training and
testing sample, it is worth to mention that in the first two experiments the parameter is within 15-20%, when in 6 eras and 6 convolutional layers the parameter is within 10%. Because of it, summing up is the proportion is optimal. The archived accuracy on the training is 84.82%. On figure 5, there are graphs of the accuracy and error dependence on an amount of iterations.

![Figure 5](image.png)

**Figure 5.** Dependency graphs of accuracy and error against the number of iterations.

The result of research is the most satisfied because while an amount of eras is increasing, accuracy of definition rises, the probability of an error reduces particularly. The difference between the percentage of an error and accuracy fluctuates within 5-10%.

7. **Results of the research**

In the way, as an final structure of neural network there was using a network with six convolutional and three sub-sampling layers which form an input vector of features for two fully connected layers. To activate there is applying ReLU function after each convolutional layers and a one fully connected layer, after the last layer there was using a non-linear sigmoid function.

The convolutional neural network converges in:

- Adam optimizer is in lr=1e-5 (the training speed).
- Size of mini-sample is 20 images.
- Six eras for training with 6 convolutional layers of neural network.

The architecture of neural network is in figure 6.
The analysis of the obtained result can display the using of the neural network with developed schema of the images pre-processing allows to say about a successful working of classifier in the recognition of ischemic stroke.

This kind of research is necessary because the earlier and more precisely ischemic stroke is detected, the more effective a therapy is.

Also, there is planning to train neural network, according actual values of MDC (measured diffusion coefficient) for full picture because the high M-sing in DWI is confirmed by pathological diffusion restrictions in MDC images [20].

The main disadvantages of the developed neural network are insufficient accuracy owing to a lack of data.

To get result that is more significant for the problem, it is necessary to enlarge an amount of data in 10 times.

8. Conclusion
To sum up, on the example of the problem there was considering the process of choosing needed architecture for a neural network and also setting of its parameters for the recognition task of objects in images.

To solve the recognition problem of ischemic stroke there was setting architecture based on VGG16, which is a deep neural network and is applying for recognition with several classes.

In the result, the neural network trained on MRI images can provide the ability to achieve the high result in the studying of a disease as ischemic stroke.

The developed neural network will allow helping doctors to make decision, reducing time of diagnostics, and removing human factor in the diagnostics.

References
[1] Convolutional neural network Infopedia for in-depth knowledge. Retrieved from: https://itnan.ru/post.php?c=1&p=348000/
[2] Milov A V, Tynchenko V S, Kukartsev V V, Tynchenko V V and Bukhtoyarov V V 2018 Use of artificial neural networks to correct non-standard errors of measuring instruments when creating integral joints Journal of Physics: Conference Series 1118(1) 012037
[3] Tynchenko V S, Tynchenko V V, Bukhtoyarov V V, Kukartsev V V, Kukartsev V A and Ereemeev D V 2019 Application of Kohonen self-organizing maps to the analysis of enterprises' employees certification results IOP Conference Series: Materials Science and Engineering 537(4) 042010
[4] Bukhtoyarov V V, Tynchenko V S, Petrovsky E A, Kukartsev V V and Kuklina A I 2018 Evolutionary method for automated design of models of vortex flowmeters transformation function Journal of Physics: Conference Series 1118(1) 012041
[5] Tynchenko V S, Petrovsky E A and Tynchenko V V 2016 The parallel genetic algorithm for
construction of technological objects neural network models 2016 2nd International Conference on Industrial Engineering. Applications and Manufacturing, ICIEAM 2016 - Proceedings 7911573

[6] Tynchenko V S, Tynchenko V V, Bukhtoyarov V V, Tynchenko S V and Petrovskyi E A 2016 The multi-objective optimization of complex objects neural network models Indian Journal of Science and Technology 9(29) 99467

[7] World Health Organization Retrieved from: https://www.who.int/ru/news-room/fact-sheets/detail/the-top-10-causes-of-death/

[8] Image resampling Online encyclopedia Wikipedia Retrieved from: https://ru.wikipedia.org/wiki/Обработка_изображений/

[9] Golovinov A O and Klimova E N 2017 The advantages of neural networks comparing to traditional algorithms The fifth intern. Scien.-practic. Conf. 5 11-5

[10] Petrov S P 2013 Convolutional neural network for recognizing car license plate characters System Analysis in Science and Education 3 66-73

[11] Ha L M 2016 Convolutional neural network for solving the classification problem Proceedings of the Moscow Institute of Physics and Technology 8(3) 91-7

[12] Bengio Y and LeCun Y 2007 Scaling learning algorithms towards Large Scale Kernel Machines 5 (Massachusetts Institute of technology press MIT Press) 1-41

[13] Lanczos resampling Online encyclopedia Wikipedia Retrieved from: https://ru.wikipedia.org/wiki/Lanczos_resampling/

[14] Filtering images in Python Infopedia for in-depth knowledge Retrieved from: https://php-academy.kiev.ua/blog/image-filtering-in-pythonD/

[15] Cherny S A 2012 Frequency and spatial methods of digital image filtering Youth Scientific and Technical Bulletin 5 31

[16] VGG16 - convolutional network for the selection of signs of images Infopedia for in-depth knowledge Retrieved from: https://neurohive.io/vidy-neurosetej/vgg16-model/

[17] Sozykin A V 2017 A review of teaching methods for deep neural networks Bulletin of the South Ural State University. Series: Computational Mathematics and Computer Science 6(3) 28-56

[18] Glorot X, Bordes A and Bengio Y 2011 Deep Sparse Rectifier Neural Networks J. of mach. learn. res. 15 315-23

[19] Susha AV 2017 Analysis of methods for optimizing the parameters of neural networks on the example of the problem of image classification J. of elect. scient. Retrieved from: https://sibac.info/journal/student/8/79708/

[20] Diffusion-weighted images (DWI) Infopedia for in-depth knowledge. Retrieved from: https://radiographia.info/article/diffuzionno-vzvshennye-izobrazheniya/