

# Convolutional neural network for detection of pathological changes in MR images of the brain

Yulia Agafonova  
Faculty of Mathematics  
Samara National Research University  
Samara, Russia  
agafonova.julia132@gmail.com

Andrey Gaidel  
Video Mining Laboratory  
Image Processing Systems Institute of  
RAS - Branch of the FSRC  
"Crystallography and Photonics" RAS  
Samara, Russia  
andrey.gaidel@gmail.com

Pavel Zelter  
Head of Radiology Department  
Clinic of the Samara State  
Medical University  
Samara, Russia  
pzelter@mail.ru

Aleksandr Kapishnikov  
Head of the Department of Radiology  
and Radiation Therapy  
Samara State Medical University  
Samara, Russia  
a.kapishnikov@gmail.com

**Abstract**—At the present day, the problem is subsisting associated with a reliable diagnosis as soon as possible, especially in medicine in cases of diagnosis of neoplasms. The article discusses research method for the diagnosis of brain diseases in magnetic-resonance tomography images, based on deep learning. This paper presents a novel approach to solutions pattern classification, was formed the optimal architecture convolutional neural network. As a result of experimental studies, was undertake a study major characteristic of convolutional neural network. Through the use of this neural network architecture 95 % the images from the validation set were classified correctly. In addition, the results can be used as an intermediate result for further images analysis.

**Keywords**—computer vision, image processing, magnetic-resonance imaging, classification, convolutional neural network

## I. INTRODUCTION

There is a problem in medicine to make a reliable diagnosis in the shortest possible time due to the growing volume of medical research [1, 2]. This problem is especially acute in cases of diagnosis of various neoplasms. The presented method can help to solve this problem. This method is able to classify a considerable number of images of the magnetic resonance imaging (MRI) of the brain into two types. The first type includes magnetic resonance imaging, in which any neoplasms are absent. The second type includes magnetic resonance imaging in which there is some neoplasms. Only magnetic resonance images of the second type demand the doctor attention. It is this difference that can accelerate the process of diagnosis of various neoplasms.

An algorithm for solving a similar problem using the classifier was presented in [3] and [4]. The described classifier was based on the use of an ensemble of decision trees [5]; as input a set consisting it inputs 98 images of the first type and 98 images of the second type. The size of all images was  $512 \times 512$  samples in size. An equal class ratio was made for maximum objectivity of the results. The above algorithm assigned each image sample to the area of pathological changes or to the background, so it actually solved the problem of classifying each sample. However, a significant drawback of this method is the voluminous preprocessing of the input images to achieve high classification results.

This article was based on the idea of achieving higher classification rates on the same set of inputs, but on the basis of a different, more modern method. To show the effectiveness of this method over an algorithm based on an ensemble of decisive trees, this study will use the following quality metrics:

– precision

$$J_P = N_{TP} / (N_{TP} + N_{FP}),$$

– recall

$$J_R = N_{TP} / (N_{TP} + N_{FN}),$$

– F1-score

$$F_1 = 2J_P J_R / (J_P + J_R),$$

– specificity

$$J_S = N_{TN} / (N_{TN} + N_{FP}),$$

– accuracy

$$J_A = (N_{TP} + N_{TN}) / (N_{TP} + N_{TN} + N_{FP} + N_{FN}), \quad (1)$$

where is  $N_{TP}$  – number of correctly classified images with pathological changes,  $N_{TN}$  – number of correctly classified images without pathological changes,  $N_{FP}$  – number of images without any pathological changes classified as they have one, and  $N_{FN}$  – number of images with pathological changes classified as they have no pathological changes.

The sources of the research materials were the archives CENTRAL, supported Washington University School of Medicine, on the platform XNAT [6] and The Cancer Imaging Archive [7], established by the Federal Research Center - Frederick National Laboratory for Cancer Research. Meningiomas and glioblastomas were present in the images. Table I shows the metrics of an algorithm based on an ensemble of decisive trees, which can be considered as a baseline. The shape features of objects extracted from the image using adaptive threshold processing. F1-score was assumed as the main metric to assess the quality of the algorithm. It is on this metric that the effectiveness of various changes to the quality of classification will be compared throughout the work.

TABLE I. THE RESULT OF THE RUNNING OF THE CLASSIFIER

Measure	MRI with neoplasms	
	$J$	<i>mistake</i>
$J_P$	0.81	0.19
$J_R$	0.97	0.03
$F_I$	<b>0.88</b>	0.12
$J_S$	0.77	0.23
$J_A$	0.87	0.13

## II. CONVOLUTIONAL NEURAL NETWORK

Currently, the common method for solving classification problems is the use of convolutional neural networks. They were first described in the last century [8]. This approach is effective for solving a wide range of tasks. However, to achieve maximum efficiency, it is necessary to take into account many factors that affect the training of neural systems [9]. In this paper, we will consider the conducted experimental researches to determine optimal convolutional neural network architecture to solve the task. The original architecture is similar in architecture to AlexNet [10].

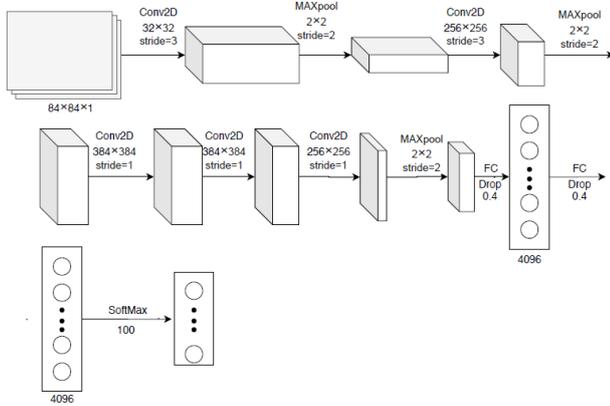


Fig 1. Convolutional neural network Architecture.

As shown in Fig. 1, initially the network consists of 5 convolutional layers, subsampling layers and fully connected layers. Input images were downscaled from  $512 \times 512$  reports to  $72 \times 72$  pixels. Training lasted for 120 epochs. On convolution and dense layers (except for the last dense layer), the ReLu activation function was used:

$$f_{\text{ReLU}}(x) = \max(x, 0).$$

To estimate losses, we used binary cross-entropy (2), where  $N$  is the number of images in the sample,  $y_i$  - is the class of the  $i$ -th image,  $p_i$  - is the output of the neural network for the  $i$ -th image.

$$H_p = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)) \quad (2)$$

Table II shows the performance metrics (1) for the convolutional neural network, which was designated as the original. Compared to Table I, the F1-score is 3% higher, however, this change does not affect the quality of classification so significantly.

To evaluate the performance indicators, we also used the ROC - curve [11] and the Precision-Recall curve [12]. The curves are shown in Figure 2.1 and Figure 2.2, respectively.

TABLE II. THE RESULT OF THE RUNNING OF THE SOURCE NETWORK

Measure	MRI with neoplasms	
	$J$	<i>mistake</i>
$J_P$	0.83	0.17
$J_R$	1.00	0.00
$F_I$	<b>0.90</b>	0.10
$J_S$	0.71	0.29
$J_A$	0.88	0.12

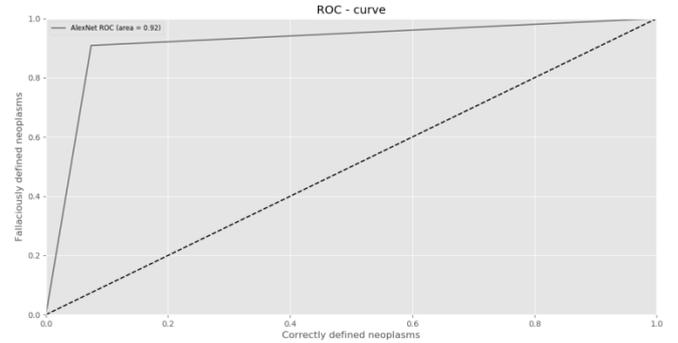


Fig. 2.1. Changes in the classification assessment parameters during the training of a convolutional neural network on the ROC graph .

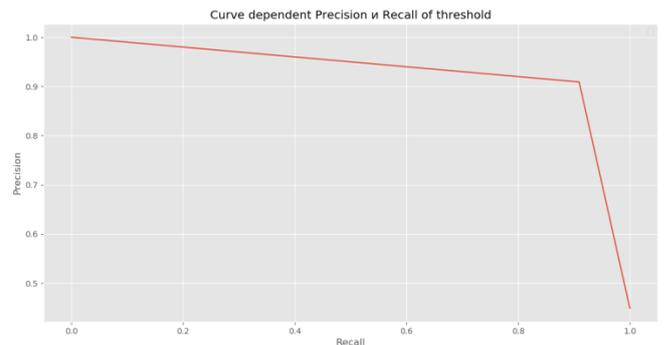


Fig. 2.2. Changes in the classification assessment parameters during the training of a convolutional neural network on the Precision-Recall curve graph.

The descriptive power of evaluating the effectiveness of classification using Precision-Recall curve was discussed in [13]. The Precision-Recall curve principle is based on mean accuracy:

$$AP = \sum_n (R_n - R_{n-1}) P_n$$

where  $R_n$  and  $P_n$  precision and recall on threshold  $n$ .

## III. RESEARCH OF EFFICIENCY OF THE CLASSIFICATION ALGORITHM BASED ON A CONVOLUTIONAL NEURAL NETWORK

### A. Effect of input image size on algorithm performance

To improve quality assessments (1), it was decided to consistently research the effect of various characteristics of the convolutional neural network architecture on the efficiency. The first research was to determine the effect of the size of the input image. As we have already said in Section 2, the original image was  $512 \times 512$  pixels, and was downscaled to  $72 \times 72$  pixels. To avoid significant changes in the speed of the convolutional neural network, the research was performed on images of the brain MRI of the

following sizes:  $32 \times 32$ ,  $40 \times 40$ ,  $48 \times 48$ ,  $56 \times 56$ ,  $64 \times 64$ ,  $72 \times 72$ ,  $84 \times 84$ ,  $92 \times 92$ . It should be emphasized the original image size remains the same,  $512 \times 512$  samples, directly resized input image.

Figure 3 shows a graph of the dependence of the size of the input image on the F1-score and accuracy. As we can see at the graph, it takes the greatest value at the sizes of the input image in  $84 \times 84$  and  $92 \times 92$  pixels. The size of  $92 \times 92$  pixels was chosen as the most favorable size for the input image in solving this problem.

The quality assessment of the algorithm with changing the size of the input image is presented in table II.

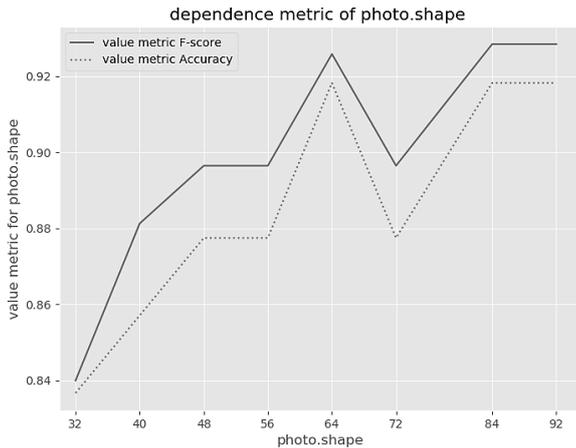


Fig 3. Dependence of the F1-score and the accuracy on changes in input image size parameters for a convolutional neural network.

TABLE III. THE RESULTS OF THE RESEARCH OF THE EFFECT OF CHANGING THE SIZE OF THE INPUT IMAGE ON THE EFFICIENCY OF THE ALGORITHM

Measure	MRI with neoplasms	
	$J$	$mistake$
$J_P$	0.89	0.17
$J_R$	0.96	0.00
$F_I$	<b>0.92</b>	0.10
$J_S$	0.86	0.29
$J_A$	0.92	0.12

### B. Effect of changes in convolutional layers on algorithm efficiency

The next stage of the work was the study of the effect of convolutional layers of the neural network on the efficiency of the algorithm.

Figure 4 shows the F1-score and the accuracy depending of the number of the study. On the x-axis the numbers of experiments (from 1 to 5) are shown. In the experiment No. 1, the activation function of the second and the third layer was replaced to sigmoid. In the experiments No. 2 and No. 3, the number of filters in the third layer was increased from 384 to 512 and the number of filters in the second layer was increased from 256 to 384, respectively. An additional sixth convolutional layer (between the fourth and fifth layer) was also added in the experiment No. 5. This layer had a convolution kernel  $3 \times 3$  and 256 filters. In the experiment No. 4, on the contrary, there were only four convolutional

layers. The layer between the third and fifth convolutional layers has been removed from the original architecture.

As shown in the Fig. 4, these experiments did not bring positive dynamics in improving the quality of classification, but rather worsened the achieved result. Moreover, in the experiment No. 5, overfitting occurred [14].

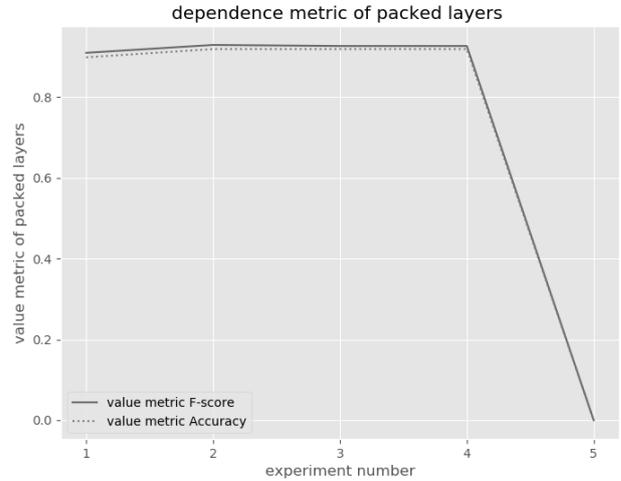


Fig 4. The effect of changes in convolutional layers on the efficiency of the algorithm.

### C. The influence of the numerical values of convolution kernels on the efficiency of the algorithm

The final stage in the work was the research of the influence of the numerical values of the convolutional kernels of the convolutional network layers on the classification results. Initially, the second layer, according to the original architecture of the convolutional neural network, had a kernel of  $11 \times 11$  samples. As experiments, convolution kernels of  $9 \times 9$ ,  $7 \times 7$ ,  $13 \times 13$  samples were taken. The results of the research are shown in table IV.

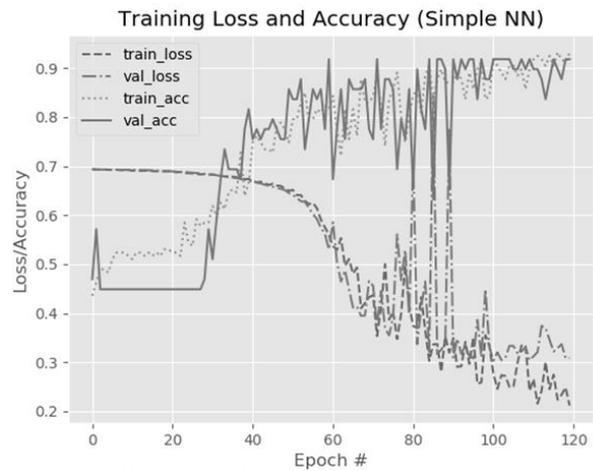


Fig 5. The effectiveness of neural network training depending on the era number (train\_acc and val\_acc are the confidence values (1) for the training and for the control sample, train\_loss and val\_loss are the values of the loss function (2) for the training and for the control sample).

By so doing, as can be seen from the table, the convolution kernel of  $9 \times 9$  samples is the most effective for solving the considered problem for this convolutional neural network architecture. To estimate the losses of the final model, binary cross-entropy (2) was used. The assessment was made after each era, this can be seen in Fig. 5.

TABLE IV. CONSEQUENCES OF CONVOLUTION KERNEL CHANGES

Measure	kernel $9 \times 9$		kernel $7 \times 7$		kernel $13 \times 13$	
	<i>J</i>	<i>mistake</i>	<i>J</i>	<i>mistake</i>	<i>J</i>	<i>mistake</i>
$J_P$	0.93	0.07	0.92	0.08	0.92	0.08
$J_R$	0.96	0.04	0.92	0.08	0.85	0.15
$F_I$	0.95	0.05	0.92	0.08	0.88	0.12
$J_S$	0.90	0.10	0.90	0.10	0.90	0.10
$J_A$	0.94	0.06	0.92	0.08	0.87	0.13

#### IV. CONCLUSION

The article examined various methods for improving the classification of brain MRI images based on a convolutional neural network. An original convolutional neural network architecture was developed that provides maximum efficiency for solving this problem.

In the course of experimental studies on a set of MRI images, the average quality metrics of classification algorithm became higher than the metrics of another method based on an ensemble of decision trees. At the same time, the need for expensive preprocessing of the input data has disappeared, which may give an increase in the speed of processing MRI images when applied in practice. The value of the F1-score was 95%, and the probability of erroneous classification was equal to 6%. This shows that the classification accuracy is higher than the method described in section 1. By so doing, we can recommend using the resulting convolutional neural network architecture for the recognition of pathologies in images of brain MRI.

These research results can be used to create a computer system for the diagnostics of various pathologies from MRI images of the human brain.

#### ACKNOWLEDGMENT

The work was partially funded by the Russian Foundation for Basic Research under grants No. 19-29-01235 and 19-29-01135 (theoretical results) and the RF Ministry of Science and Higher Education within the government project of the FSRC "Crystallography and Photonics" RAS under grant No. 007-GZ/Ch3363/26 (numerical calculations).

#### REFERENCES

- [1] N.A. Smelkina, R.N. Kosarev, A.V. Nikonorov, I.M. Bairikov, K.N. Ryabov, A.V. Avdeev and N.L. Kazanskiy, "Reconstruction of anatomical structures using statistical shape modeling," *Computer Optics*, vol. 41, no. 6, pp. 897-904, 2017. DOI: 10.18287/2412-6179-2017-41-6-897-904.
- [2] N.Yu. Ilyasova, N.S. Demin, A.S. Shirokanov, A.V. Kupriyanov and E.A. Zamytskiy, "Method for selection macular edema region using optical coherence tomography data," *Computer Optics*, vol. 44, no. 2, pp. 250-258, 2020. DOI: 10.18287/2412-6179-CO-691.
- [3] J.D. Agafonova and A.V. Gaidel, "Localization of the area of pathological changes in the images of brain MRIs," Samara: Advanced Information Technology (AIT), pp. 362-365, 2019.
- [4] Yu.D. Agafonova, A.V. Gaidel, P.M. Zelter and A.V. Kaphishnikov, "Efficiency of machine learning algorithms and convolutional neural network for detection of pathological changes in MR images of the brain," *Computer Optics*, vol. 44, no. 2, pp. 266-273, 2020. DOI: 10.18287/2412-6179-CO-671.
- [5] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [6] R. Herrick, W. Horton, T. Olsen, M. McKay, K.A. Archie and D.S. Marcus, "XNAT Central: Open sourcing imaging research data," *Neuroimage*, vol. 124, no. B, pp. 1093-1096, 2016.
- [7] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox and F. Prior, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045-1057, 2013.
- [8] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [9] K. He and J. Sun, "Convolutional neural networks at constrained time cost," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5353-5360, 2015.
- [10] A. Krizhevsky, I. Sutskever and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017.
- [11] T. Fawcett, "An introduction to ROC analysis," *J. Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006.
- [12] J. Davis and M. Goadrich, "The Relationship Between Precision-Recall and ROC Curves," *ICML*, 2006.
- [13] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS ONE*, vol. 10, no. 3, e0118432, 2015.
- [14] V.N. Vapnik, "Statistical learning theory," N.Y.: John Wiley & Sons, Inc., 1998.