

CONVOLUTIONAL NEURAL NETWORKS IN THE TASK OF IMAGE CLASSIFICATION

**Assoc. Prof. Larisa Zelenina,
Assoc. Prof. Liudmila Khaimina, Evgenii Khaimin,
D. Khripunov, Assoc. Prof. Inga Zashikhina**

Northern (Arctic) Federal University named after M.V. Lomonosova (Russia)

Abstract. Convolutional neural networks are acquiring general acknowledgement for diverse application areas. The article describes the process of solving the task of images classification using convolutional neural networks. The authors present the examples of using convolutional neural networks for various purposes. The composed set of data is used to implement and train the model of convolutional neural network for the task of classification of medical images.

Keywords: convolutional neural network; image classification problem; Python programming language; machine learning

Introduction

Classification problem in machine learning has been successfully used for the character recognition tasks (Kalaichelvi & Ahammed Shamir Ali 2012; Zelenina, Khaimina, Khaimin, Antufiev & Zashikhina 2020; Zelenina, Khaimina, Khaimin, Khripunov & Zashikhina 2021). The solution to the classification problem is to implement by machine learning methods the algorithm that allows to identify whether an object belongs to a certain class of objects, into which the original set is divided and there is a set of objects with known class membership.

Classification can be binary (the dependent variable takes only two values (yes/no or 0/1) or multiple (the dependent variable takes values from some set of classes) classification. The estimation of classification accuracy is done by cross-validation on the basis of comparison of accuracy on the training and test sets.

One of the methods of solving classification problems is based on using artificial neural networks, namely convolutional neural networks (CNN). Their wide application is due to the rather high accuracy and speed of the solution in the tasks of searching objects. CNNs provide partial stability to the following types of image distortions: displacements, rotations, changes of scale, changes of perspectives and

so on. In this paper, we are going to demonstrate the use of CNN for the task of classification of medical images.

Methodology

To measure the accuracy of the designed convolutional neural network, a loss function is used. This function returns a real number reflecting the accuracy of the neural network response. The Categorical Cross-Entropy Loss Function can be used. Its formula is shown below:

$$E = - \sum_{i=0}^n y_i^{truth} \ln(y_i^l),$$

where y_i^l is the answer received at the output of the network, y_i^{truth} is the correct answer, n is the number of classes.

Neural network training is reduced to the minimization of the loss function by adjusting the weights of synaptic connections between neurons. The algorithm of backpropagation (or generalized delta rule) is used for that.

Convolutional neural networks in the task of image classification

The current opportunities of applying convolutional neural networks in image processing are extensive: they can localize objects, mark up an image by singling out its components. The CNN algorithms can determine the sex and age of a person in the image by identifying his/her emotions. These technologies are now becoming common. In addition, convolutional neural networks can allow to “transfer” the style of one image to the content of another image, obtaining absolutely new images. In this case the new images are based on the image that sets the content and the image that sets the features. It is possible to train a neural network to read lips”, i.e. on the basis of the image of a speaking person, the neural network will allow to map the image of a speaking person into the text of their conversation.

The use of CNN in the task of image classification is a highly relevant topic of scientific research and has a wide range of practical applications.

The use of deep learning methods can be considered an alternative methodology for the detection of extreme climatic phenomena. For example, CNNs can be used to solve the problems of detection of climatic patterns (Liu, Racah, Prabhat, Correa, Khosrowshahi, Lavers, Kunkel, Wehner & Collins 2016). Images of climate phenomena such as tropical cyclones, atmospheric rivers, and atmospheric fronts were used to train the networks. The proposed CNN architecture for solving climate pattern recognition problems, combined with a Bayesian hyperparameter optimization scheme, achieved 89 – 99% accuracy in the detection of extreme phenomena.

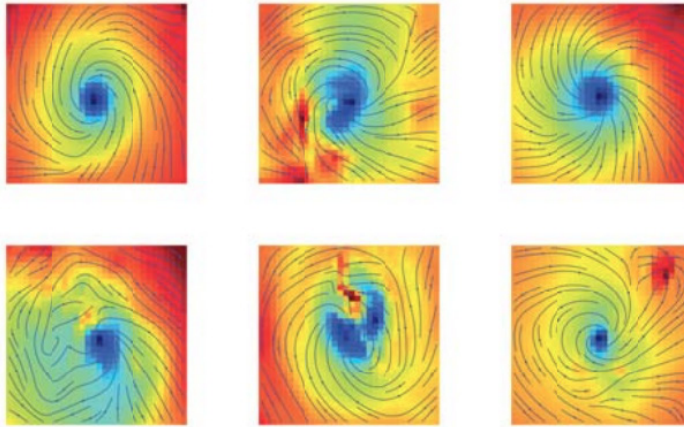


Figure 1. Examples of images of tropical cyclones

Pattern recognition systems based on computer vision technology are tested in various industries. An illustration is a model that identifies a car license plate taking into account all the irregularities and different fonts, including non-Latin fonts (Kirad Varad Vinay, Indla Omkar Balaobaiah, Mujawar Sohail Mahiboob, & Shin-de Dinesh Nagnath 2021). To identify characters from a license plate, optical character recognition (OCR) technology is used, which consists of two parts: character segmentation and character recognition.



Figure 2. Example of identified number from the image

The developed software highlights the license plate in the received image and then segments it into individual characters for license plate identification.

Automated vehicle tracking systems are used to track fast-moving vehicles with roadside cameras. The CNN You Only Look Once (YOLO) architecture, which is used to recognize multiple objects in an image, is used to detect objects.

In this process, video footage is first converted into images to identify a vehicle using a trained convolutional neural network (CNN). Then the license plate number is determined, also with the help of a CNN working on sign recognition. The last step is optical character recognition (OCR), which is used to read the characters from the previously detected license plate. The proposed model uses the ImageAI library to simplify the learning process (Gnanaprakash, Kanthimathi & Saranya 2021).

Another challenge for CNN could be the face recognition problem. The proposed 4-layer CNN architecture is capable of processing face images that contain occlusions, poses, facial expressions, and variable lighting.



Figure 3. Examples of images from the dataset

The training of the built model took place on a sufficiently large dataset. Nevertheless, the proposed system completed the face recognition process in less than 0.01 seconds (Syafeeza, Khalil-Hani, Liew & Bakhteri 2014).

If the initial dataset is small, then the solution to the problem of human face recognition can be based on an approach that combines a CNN with augmentation of the dataset. In this case, the original small dataset is augmented to a large dataset through several transformations of face images. Then, based on the augmented dataset, facial features can be efficiently extracted, and recognition accuracy can be improved using the developed CNN (Lu, Song, & Xu 2020).

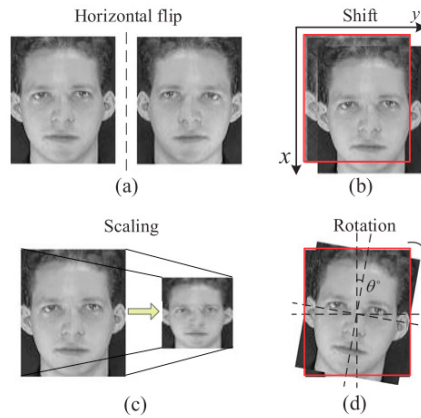


Figure 4. Example of image transformation for data augmentation

Current trends in deep learning applied to medicine are intensifying. Convergent neural networks are widely used for classification and segmentation of medical images, allowing to accelerate the process of diagnosis. For example, a trained CNN based on a marked dataset of medical images of infected and uninfected lungs (obtained from CT scans), will then be able to classify a new image as confirming or not confirming the disease (Kugunavar & Prabhakar 2021).

Relevant at this time is an analysis of the current state of CNN models for COVID-19 detection based on images obtained by radiological diagnostic methods.

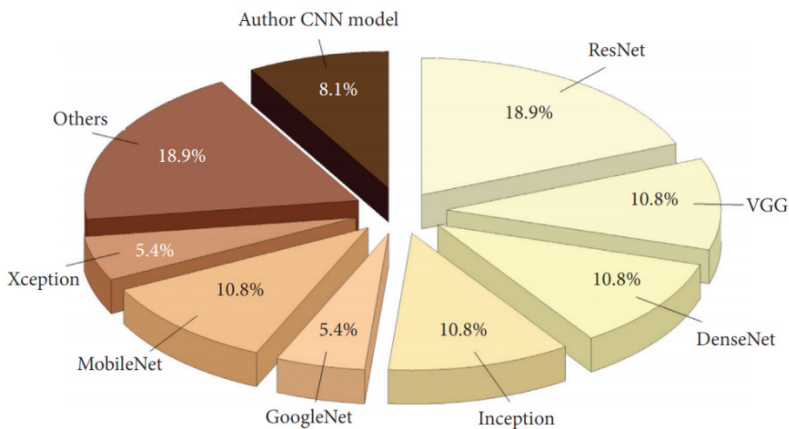


Figure 5. Frequency of using different CNN architectures when processing radiological images

In the nearest future, it will be possible to achieve faster, cheaper and safer disease diagnosis based on the use of deep learning algorithms in radiology centers (Ghaderzadeh & Asadi 2021).

Modeling a neural network for the analysis of medical images

Let us consider the process of modeling a neural network for X-ray scans processing.

Neural network development can be based on libraries (TensorFlow, Fastai, Keras) and modules (shutil and os) of Python programming language.

To train the neural network we used the dataset (high resolution X-ray images of lungs with a confirmed diagnosis) of the opensource site kaggle.com. We applied about 25% of the dataset as for the validation sample, about 15% – for the test sample. The ImageDataGenerator method of the Keras library made it possible to generate individual samples, creating datasets based on the available images, as well as their augmentation.

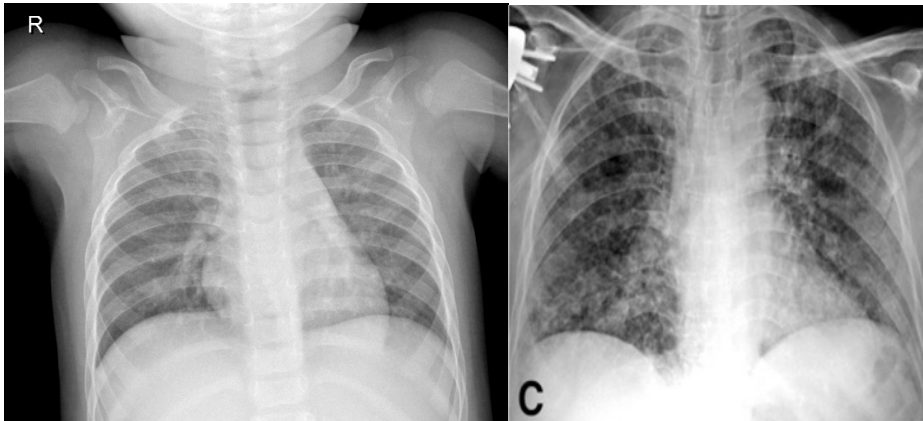


Figure 6. Examples of dataset images

Below we present the listing fragment of the convolutional neural network simulation:

```
model.add(Conv2D(16, (5, 5), activation = 'relu', input_shape
= input_shape))
model.add(MaxPooling2D(pool_size = (2, 2)))
model.add(Dropout(0.2))

model.add(Conv2D(32, (5, 5), activation = 'relu'))
```

```
model.add(MaxPooling2D(pool_size = (2, 2)))
model.add(Dropout(0.2))

model.add(Conv2D(64, (5, 5), activation = 'relu'))
...
...
model.add(Dense(2, activation = 'softmax'))
```

A sequential-type model was used during construction. The convolution layer Conv2D contains:

- matrix of filters (5x5 dimensionality)
- 16 filters
- ReLU activation function
- input_shape = (512, 512, 3) (image dimension 512x51, 3 rgb channels).

The MaxPooling2D sub-sample layer is a layer of maximal pooling with reduction to 1 pixel of a group of 2x2 pixels.

Dropout is used to prevent overtraining of the network during training. It consists in selective disconnection of neurons during training. So, when some neurons are switched off during the learning process, other neurons are more active. The quality of network learning also increases. Overall, the trained network becomes more stable. The probability of exclusion of each neuron at any training iteration is determined.

The content of subsequent blocks of the model is similar, with the content in the convolution layers varying from 32 to 64.

The Flatten layer converts the resulting data into a vector and is located between the convolutional layers and the full-link layer.

The Dense layer contains 128 neurons and a ReLU activation function.

The Dropout layer consists of neurons, the number of which is equal to the number of classes. In order to be able to interpret the obtained result as a number of possible outcomes – Softmax (activation function) is used. It is employed to convert the obtained sum to one.

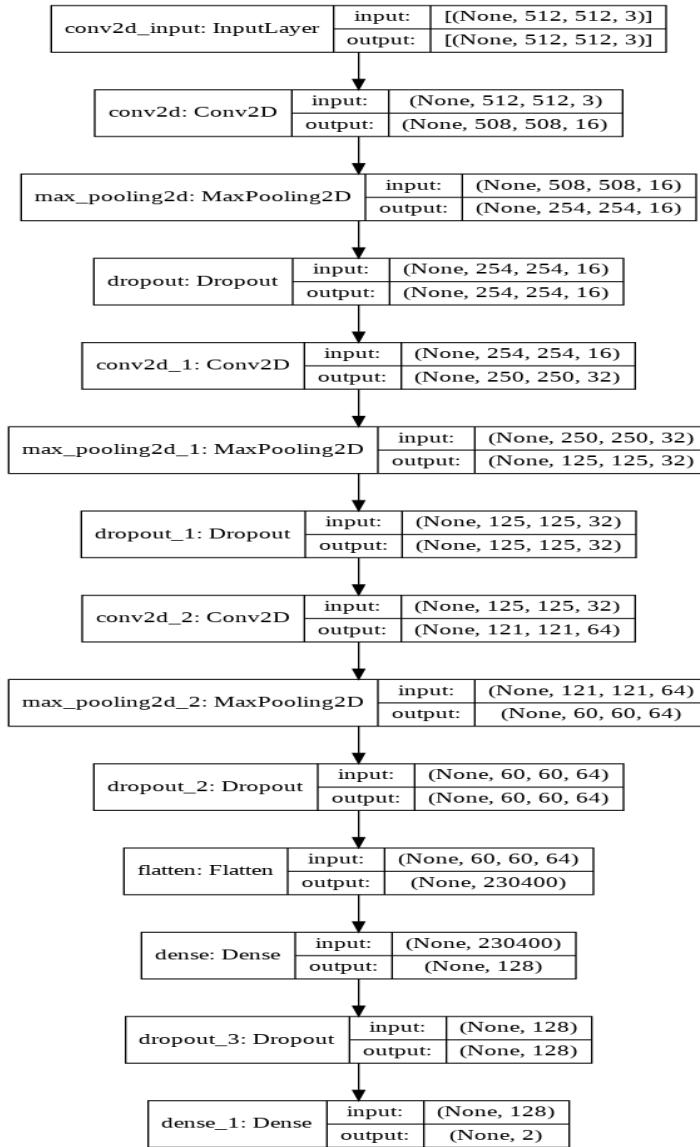


Figure 7. Fragment of the neural network architecture diagram

Training of the compiled model was performed over 20 epochs. The accuracy of the model on the dataset test data was 95%. Figure 3 presents the error dependence on the number of epochs of network training on the first 10 epochs.

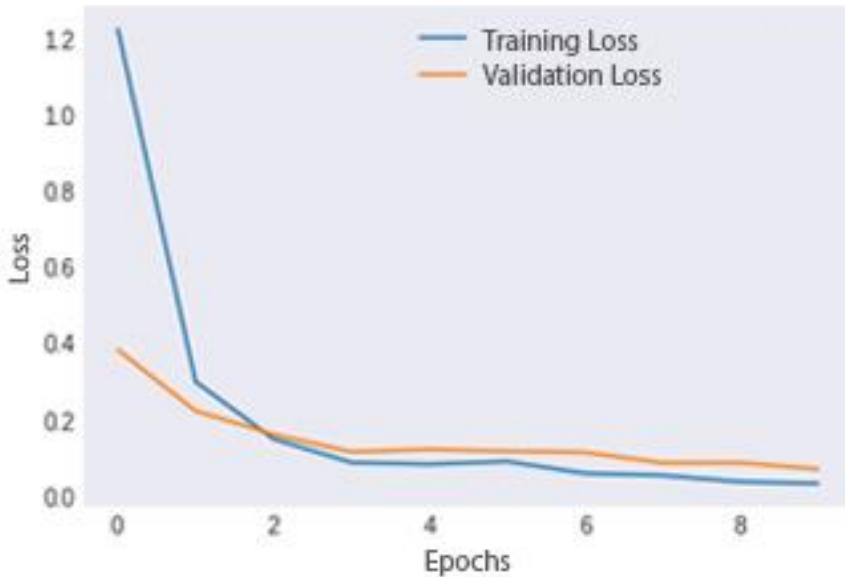


Figure 8. Accuracy and error of the neural network

Conclusion

Convergent neural networks (CNNs), which are a class of deep learning neural networks, enable image recognition and classification. The task of image classification is of great value. This paper deals with CNN modeling for the diagnosis of lung disease using X-ray images. An open-source dataset representing high-resolution X-ray images of the lungs with a confirmed diagnosis was used to train the constructed network. Analysis of the simulated neural network accuracy indicates its ability to confirm/reject a patient's diagnosis based on the submitted medical images.

Our research was aimed at the development of an application for the efficient and rapid detection of lung disease from medical images. The essence of the work is that according to the available data of the lung disease make changes in the X-ray images of the patient. In this case, it is possible that small changes in the images are not noticed during the visual examination. Artificial Intelligence methods can help to detect the disease in a short time and make a clinical diagnosis. This saves time for disease control as well as a more accurate and earlier diagnosis of the disease. Further research is related to the discovery of new ways to use convolutional neural networks in medicine.

REFERENCES

- KALAICHELVI, V. & AHAMMED SHAMIR ALI, 2012. Application of Neural Networks in Character Recognition. *International Journal of Computer Applications* (0975 – 8887), **52**(12).
- ZELENINA, L.I., KHAIMINA, L.E., KHAIMIN, E.S., ANTUFIEV, D.I. & ZASHIKHINA, I.M., 2020. Neural Networks in a Character Recognition Mobile Application. *Mathematics and Informatics*, **63**(5), 484 – 500
- ZELENINA, L.I., KHAIMINA, L.E., KHAIMIN, E.S., KHRIPUNOV, D.D. & ZASHIKHINA, I.M., 2021. The problem of images' classification: neural networks. *Mathematics and Informatics*, **64**(3), 289 – 300.
- LIU, Y., RACAH, E., PRABHAT, CORREA, J., KHOSROWSHAHI, A., LAVERS, D.A., KUNKEL, K.E., WEHNER, M.F. & COLLINS, W.D., 2016. Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets. ArXiv, abs/1605.01156.
- KIRAD VARAD VINAY, INDLA OMKAR BALAOBAIAH, MUJAWAR SOHAIL MAHIBOOB & SHINDE DINESH NAGNATH, 2021. Automatic number plate recognition for different fonts and non-roman script. *International Journal of Engineering Applied Sciences and Technology*, **5**(11). DOI: 10.15680/IJIRCCE.2021.0906112.
- GNANAPRAKASH, V., KANTHIMATHI, N. & SARANYA, N., 2021. Automatic number plate recognition using deep learning. *IOP Conference Series: Materials Science and Engineering*, **1084**(1), 012027. doi:10.1088/1757-899x/1084/1/012027.
- SYAFEEZA, A.R., KHALIL-HANI, M., LIEW, S.S. & BAKHTERI, R., 2014. Convolutional Neural Network for Face Recognition with Pose and Illumination Variation. *International journal of engineering and technology*, **6**, 44 – 57.
- LU, P., SONG, B. & XU, L., 2020. Human face recognition based on convolutional neural network and augmented dataset. *Systems Science & Control Engineering*, **9**(sup 2), 29 – 37. doi:10.1080/21642583.2020.1836526.
- KUGUNAVAR, S. & PRABHAKAR, C. J., 2021. Convolutional neural networks for the diagnosis and prognosis of the coronavirus disease pandemic. *Visual computing for industry, biomedicine, and art*, **4**(1), 12. <https://doi.org/10.1186/s42492-021-00078-w>.
- GHADERZADEH, M. & ASADI, F., 2021. Deep Learning in the Detection and Diagnosis of COVID-19 Using Radiology Modalities: A Systematic Review. *Journal of healthcare engineering*, 6677314. <https://doi.org/10.1155/2021/6677314>.

✉ **Dr. Larisa I. Zelenina, Assoc. Prof.**

ORCID iD: 0000-0002-0155-3139

Department of Applied Mathematics and High-performance Computing
Higher School of Information Technologies and Automated Systems
Northern (Arctic) Federal University named after M.V. Lomonosov
Arkhangelsk, Russia
E-mail: l.zelenina@narfu.ru

✉ **Dr. Liudmila E. Khaimina, Assoc. Prof.**

ORCID iD: 0000-0003-4552-0440

Department of Applied Informatics and Information Security
Higher school of Information Technologies and Automated Systems
Northern Arctic Federal University named after M.V. Lomonosov
Arkhangelsk, Russia
E-mail: l.khaimina@narfu.ru

✉ **Evgenii S. Khaimin, Senior Lecturer**

ORCID iD: 0000-0003-0523-3623

Department of Applied Informatics and Information Security
Higher school of Information Technologies and Automated Systems
Northern Arctic Federal University named after M.V. Lomonosov
Arkhangelsk, Russia
E-mail: e.khaymin@narfu.ru

✉ **Mr. D.S. Khripunov, Master Student**

Department of Applied Mathematics and High-performance Computing
Higher school of Information Technologies and Automated Systems
Northern Arctic Federal University named after M.V. Lomonosov
Arkhangelsk, Russia
E-mail: hripunov_d@edu.narfu.ru

✉ **Dr. Inga M. Zashikhina, Assoc. Prof.**

ORCID iD: 0000-0002-8217-2302

Department of Philosophy and Sociology
Higher School of Social Sciences, Humanities and International Communication
Northern (Arctic) Federal University named after M.V. Lomonosov
Arkhangelsk, Russia
E-mail: i.zashikhina@narfu.ru