Impact of augmentation techniques on the classification of medical images

Antoni Jaszcz

Faculty of Applied Mathematics, Silesian University of Technology, Kaszubska 23, 44100 Gliwice, Poland

Abstract
The analysis of medical data is an important task as it can help in the quick diagnosis of the patient. This work focuses on the analysis of X-ray images. The images show the patient's condition, who is healthy or suspected of having pneumonia. To enable the automatic analysis of such images, I suggest using a convolutional neural network based on various augmentation methods. The introduction of augmentation allowed to increase the training set for the neural network, which requires a large amount of data in order to best adapt the model to the problem. The network has been described, implemented and tested to validate its operation. The research focused on various augmentation techniques including random rotation, random contrast, and a combination of both these methods. Based on obtained results, contrast augmentation achieves better results concerning the lack of its use. For the other two augmentation results, the results were lowered due to the modification of the basic orientation in the x-rays.

Keywords
Data classification, convolutional neural networks, medical images, augmentation

1. Introduction

Artificial intelligence methods allow for quick segmentation or classification of various data. However, these methods require an enormous amount of data to train such models. This is especially visible in the case of artificial neural networks, where deep architectures can classify data much better, although they need a lot of training data. Quite often, such data may not be enough to obtain a solution that can be implemented in practice. For this purpose, augmentation is used. It is the process of artificially creating new samples within a single class to generate new samples that can increase the amount of data in the training set [1].

In the case of image processing, the augmentation is based on rotating or zooming some areas. This can provide a new sample with similar features but in different orientations or configurations. Apart from the classic methods of sample analysis, new ones are proposed. An example of this is augmentation based on combining two samples based on interpolation of mathematical functions [2]. The idea is to create points from two images and interpolate them to superimpose two images with a certain transparency. Similar tools (like interpolation techniques) can be used in different approaches. It was shown in [3], where the authors use it to generate synthetic data instances. Again in [4], the idea of random cropping as an augmentation method was shown as not the best approach. According to the presented results, this method can produce noise in the gradient during the training process.

The augmentation process is very important in tasks where the data are gathered for a long time, like medicine. Automatic analysis of test results in the form of expert systems is very necessary to reduce the waiting time for a diagnosis. For this purpose, expert systems quite often use solutions based on convolutional neural networks (CNNs). It is visible in the images of moles on the skin, which are one of the basic and first examinations for the detection of potential skin cancer like melanoma. CNN's can be used for image processing, feature extraction and even classification or segmentation what was shown in [5, 6, 7]. This type of machine learning technique is also used in the detection of Parkinson's disease [8]. Medical analysis by the use of machine learning is badly needed for faster disease detection and choice of treatment. Biomedical informatics uses also augmented reality for increasing the quality of data processing and learning [9, 10, 11].

Decision support systems quite often rely not only on algorithms, but also frameworks and alternative solutions. An example of a framework for the analysis of medical images, especially those obtained during tomography, is presented in [12]. In addition, new neural network architectures are also modeled to diagnose e.g. covid-19 [13, 14]. Moreover, medical systems rely on deep neural networks that require training. The classical approach is based on teaching one model, but federal learning is also developed. it is based on training in parallel on many clients who aggregate a common model [15, 16].

Based on this observation, in this paper, a deep learning method was used to fast analysis of x-ray images to detect possible pneumonia. The contribution of this

research are:

- analysis of selected augmentation methods and its impact on convolutional neural networks,
- the use of augmentation methods to expand the training set of medical images.

2. Methodology

In this section, all mathematical aspects of CNNs, training algorithms and augmentation methods are described.

2.1. Convolutional neural network

CNN is created based on three types of layers: convolutional, pooling and fully connected (dense). The first type is the convolutional one, which has the purpose of changing the image to extract the features from the analyzed image. It is done by applying the convolutional operator($*$) on each pixel at a given position in image $I_{x,y}$ and filter matrix $k_{p \times p}$ according to:

$$k * I_{x,y} = \sum_{i=1}^{p} \sum_{j=1}^{p} k_{i,j} \cdot I_{x+i-1, y+j-1} + b, \quad (1)$$

where b is a bias.

The second layer is called pooling. The main task of it to resize the image. It is performed by the selection of one pixel from a given grid by the use of a mathematical function like minimum or maximum. After finding a pixel in the first grid (that is placed above pixel on position (0,0) in the image), the grid is moved to the next pixel. This is done until the grid does not cover the last pixel in the image. As a result of the layer's operation, an image is created from selected pixels.

The last layer is fully-connected and it is a classic column of neurons that numerical values and weights:

$$f \left(\sum_{i=0}^{n-1} w_{i,j} x_i \right), \quad (2)$$

where n is the number of neurons in the previous column, x_i is the results from a neuron in the previous layer on $w_{i,j}$ connection.

2.2. Training algorithm

Training process of CNN consists in modifying the weight values, which can be done with the ADAM algorithm [17]. This algorithm assumes that the weights will be changed according to statistical values including mean m and v in t-th iteration. The formulation of this can be defined as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (3)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2. \quad (4)$$

In the above formulas, the coefficients β_1 and β_2 are a distributions. having these two parameters, the correlation of them is calculated:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad (5)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}. \quad (6)$$

Finally, the weight in the next iteration $(t + 1)$ will be defined by the following formula:

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t, \quad (7)$$

where $\epsilon \approx 0$, and η is known as learning coefficient.

2.3. Augmentation methods

2.3.1. Random rotation

This model adds an augmentation layer that slightly and randomly rotates the input image, right before the input layer of the base model. The mathematical formulation of this method can be shown as a transformation matrix:

$$\begin{bmatrix} \alpha & \beta & (1-\alpha) \cdot x - \beta \cdot y \\ -\beta & \alpha & \beta \cdot x + (1-\alpha) \cdot y \end{bmatrix}, \quad (8)$$

where $\alpha = a \cdot \cos\theta$, $\beta = a \cdot \sin\theta$, θ is the rotation angle chosen in random way and a is a scale parameter. An example of such augmentation is shown in Fig. 1.
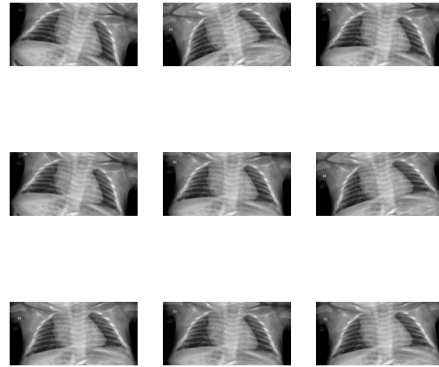

Figure 1: An example of how images are randomly rotated

2.3.2. Random contrast

This model adds an augmentation layer that slightly and randomly changes the contrast of the input image, right before the input layer of the base model. An example

of such contrast changing is presented in Fig. 2. This is made by changing the value of colors as:

$$R' = F(R - 128) + 128, \tag{9}$$

where R' is a new color in RGB color model, R is a changed value of selected color, and F is the correlation coefficient defined as follows:

$$F = \frac{259(C + 255)}{255(259 - C)}, \tag{10}$$

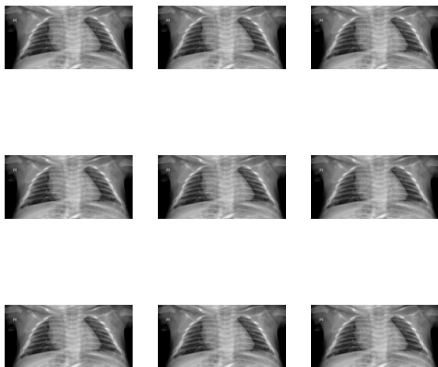where C is the contrast level. In the case of augmentation, this coefficient is random.

Figure 2: An example of how random contrast is applied to an image

2.3.3. Random rotation and contrast

This model joins two previously described augmentation methods, and applies them to the input image, right before the input layer of the base model. The combination of both presented augmentation methods is shown in Fig. 3.

3. Experiments

In this section, the experimental settings, obtained results and discussion are presented.

3.1. Testing environment

All experiments were conducted on a computer with the following specifications:
Processor: AMD Ryzen 5 5600X 6-Core Processor 4.20 GHz
Installed RAM: 32.0 GB
System type: 64-bit Windows 10 ; x64-based processor

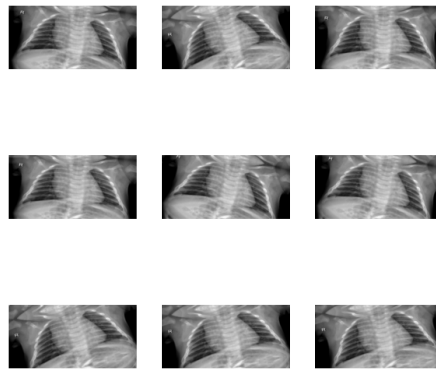All computing was done solely with the CPU.

Figure 3: An example of how random contrast and rotation is applied to an image

3.2. Database

The data used in our experiments consists of 5216 x-ray images (of different sizes) of patients with suspected pneumonia, 3875 of which were confirmed cases (viral and bacterial infections both wise), while the other 1341 were healthy. The data is accessible at *Kraggle*, at this link. *Kraggle* is a public dataset platform for data scientists and machine learning enthusiasts, controlled by *Google LLC*.

3.3. Data preparation

The images were first resized to 256x128 (pixels) and then divided randomly into two groups:

- train group (75% of the database)
- validation group (25% of the database)

3.4. Assessment

The goal of this paper is to show what impact different types of data augmentation have on an already well-performing neural network model. The structure of the base CNN model is as follows:

1. *Input layer* - a convolutional layer, having 128 neurons with 3x3-sized filters, with input shape: 128,256,1 (shape of a 2D image), with *ReLU* (Rectified Linear Unit) activation function. The output of this layer is then passed onto pooling layers, described below.

2. *Hidden layer* - in our model, we used two more convolutional blocks, with each subsequent having half the number of neurons than the previous one. All pooling layers used in the model have 2x2-sized filters. All convolutional layers used in the model have 3x3-sized filters. The output

is then passed onto the dense layer, with 64 neurons and *ReLU* activation. Next, the dropout layer (with the rate set to 0.5) and the following flatten layer prepare the final output, of the hidden segment. The order of these layers and the number of neurons within them is displayed below:

- pooling layer (2x2), *ReLU*
- convolutional layer (3x3), 64 neurons, *ReLU*
- pooling layer (2x2), activation: *ReLU*
- convolutional layer (3x3), 32 neurons, *ReLU*
- pooling layer (2x2), *ReLU*
- dense layer, 64 neurons, *ReLU*
- dropout layer (threshold: 0.5)
- flatten layer

3. *Output layer* - a dense layer with 2 neurons, related to the healthiness of a patient. While assessing, the larger value of two neurons is chosen and thus, the patient is determined as healthy or ill with pneumonia.

3.5. Results

In this subsection, results of the experiments are shown. In the Tab. 1, calculated metrics, such as *accuracy*, *precision*, *recall* and *f1 − score* are displayed for each model. The values were calculated with following formulas (for binary classification):

- Accuracy:

$$\alpha = \frac{TP + TN}{TP + FP + TN + FN}, \quad (11)$$

- Precision:

$$\psi = \frac{TP}{TP + FP}, \quad (12)$$

- Recall:

$$\rho = \frac{TP}{TP + FN}, \quad (13)$$

- F1-Score:

$$\frac{1}{f1} = 0.5 \cdot \left(\frac{1}{\psi} + \frac{1}{\rho} \right), \quad (14)$$

where
TP - true sample predicted as true,
TN - false sample predicted as false,
FP - false sample predicted as true,
FN - true sample predicted as false.
In the case of the metrics other than accuracy itself, two cases are considered. First, where pneumonia is considered as the truth and healthy as falsity. Second, where it's the other way round.
The relationship between predicted and real outcome is also displayed in confusion matrices (Fig. 4), for each tested model.

The results were obtained by assessing aforementioned validation group, in which 977 (roughly 75% of the group) cases were pneumonic, while the rest 327 (25%) were healthy.

Table 1
Calculated metrics

	class	precision	recall	f1-score
base CNN model	0	0.9630	0.8746	0.9167
	1	0.9593	0.9887	0.9738
	accuracy=0.9601			
rotation	0	0.9273	0.8196	0.8701
	1	0.9419	0.9785	0.9598
	accuracy=0.9387			
contrast	0	0.9199	0.9480	0.9337
	1	0.9824	0.9724	0.9774
	accuracy=0.9663			
rotation and contrast	0	0.8043	0.9052	0.8518
	1	0.9669	0.9263	0.9462
	accuracy=0.9210			

In Tab. 1, the results show, that the base model reaches a decent accuracy of 96%, but its recall could be improved when it comes to detecting healthy cases (41 patients out of 327 healthy ones were diagnosed with pneumonia by the model, while in fact, being healthy). Furthermore, rotation augmentation worsen the overall performance of the model, while contrast one proved to be somewhat beneficial to the results (see Fig. 4). Not only did accuracy slightly improve, but its recall metric in detecting healthy cases grew considerably. Although the ability to detect all the pneumonic cases dropped, contrast augmentation brings in more balance to the model's assessment and thus improves its overall performance.

4. Conclusions

The analysis of medical images is important in order to quickly detect or help a doctor make a diagnosis decision. For this purpose, the use of a convolutional neural network for the analysis of X-ray images was presented. As part of the research, the possibilities of using augmentation were considered (techniques such as random rotation, contrast change and a combination of both). The obtained results indicate that augmentation can quickly and easily extend the training set. Random contrast change as the main augmentation technique performed better in terms of model accuracy compared to the original database. In addition, it was found that the use of rotation on medical images deteriorated the performance of the trained model. The reason for this is the rearrangement of the chest area on X-rays. As a result, the database is enlarged with data that drastically differ from the rest, and consequently, reduces the effectiveness of the neural
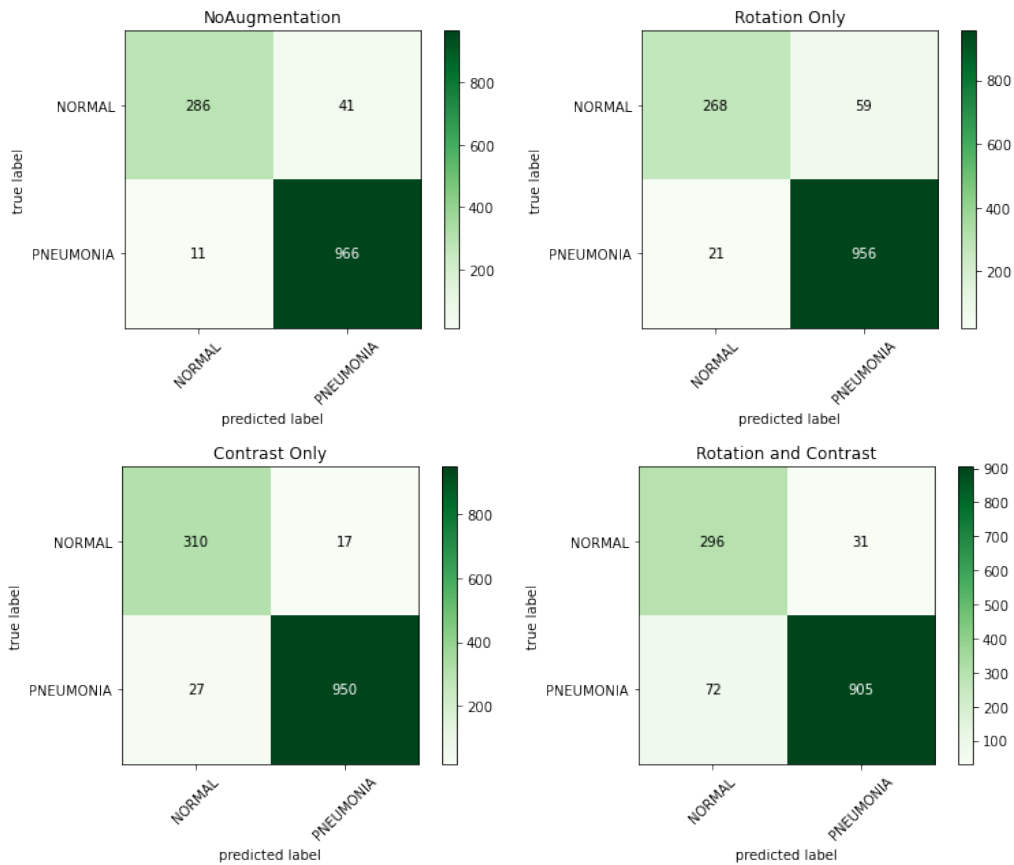
Figure 4: Confusion matrices for described models

network. The results of the last model analyzed in this paper, that is the one with both augmentations applied, show the worst results of them all. Not only its accuracy is lower, but also the ability to detect pneumonic cases, which is crucial in medical illness detection, plummeted. A positive impact of classic data augmentation techniques on CNN-model performance was similarly shown in liver illness recognition [18]. It was also suggested, that classic augmentation methods connected with cutting edge augmentation methods, such as generative adversarial network (GAN), yield the best results of all model configurations t ested. In [19], possible negative effects of joined classic augmentation methods in medical image classification were discussed, as well as their lone impact on the learning process.

Acknowledgements

References

[1] M. Elgendi, M. U. Nasir, Q. Tang, D. Smith, J.-P. Grenier, C. Batte, B. Spieler, W. D. Leslie, C. Menon, R. R. Fletcher, et al., The effectiveness of image augmentation in deep learning networks for detecting covid-19: A geometric transformation perspective, Frontiers in Medicine 8 (2021).

[2] D. Połap, M. Włodarczyk-Sielicka, Interpolation merge as augmentation technique in the problem of ship classification, in: 2020 15th Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2020, pp. 443–446.

[3] O. O. Abayomi-Alli, R. Damasevicius, S. Misra, R. Maskeliunas, A. Abayomi-Alli, Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold, Turkish Journal of Electrical Engineering & Computer Sciences 29 (2021) 2600–2614.

[4] R. Yang, R. Wang, Y. Deng, X. Jia, H. Zhang, Re-

thinking the random cropping data augmentation method used in the training of cnn-based sar image ship detector, Remote Sensing 13 (2021) 34.

[5] D. Połap, Analysis of skin marks through the use of intelligent things, IEEE Access 7 (2019) 149355–149363.

[6] R. Wang, G. Zheng, Cycmis: Cycle-consistent cross-domain medical image segmentation via diverse image augmentation, Medical Image Analysis 76 (2022) 102328.

[7] D. Połap, Fuzzy consensus with federated learning method in medical systems, IEEE Access 9 (2021) 150383–150392.

[8] O. O. Abayomi-Alli, R. Damaševičius, R. Maskeliūnas, A. Abayomi-Alli, Bilstm with data augmentation using interpolation methods to improve early detection of parkinson disease, in: 2020 15th Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2020, pp. 371–380.

[9] Y. Djenouri, A. Belhadi, G. Srivastava, J. C.-W. Lin, Secure collaborative augmented reality framework for biomedical informatics, IEEE Journal of Biomedical and Health Informatics (2021).

[10] C. Moro, J. Birt, Z. Stromberga, C. Phelps, J. Clark, P. Glasziou, A. M. Scott, Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: A systematic review and meta-analysis, Anatomical sciences education 14 (2021) 368–376.

[11] Y. Zhuang, J. Sun, J. Liu, Diagnosis of chronic kidney disease by three-dimensional contrast-enhanced ultrasound combined with augmented reality medical technology, Journal of Healthcare Engineering 2021 (2021).

[12] T. Akram, M. Attique, S. Gul, A. Shahzad, M. Altaf, S. Naqvi, R. Damaševičius, R. Maskeliūnas, A novel framework for rapid diagnosis of covid-19 on computed tomography scans, Pattern analysis and applications 24 (2021) 951–964.

[13] J. Rasheed, A. A. Hameed, C. Djeddi, A. Jamil, F. Al-Turjman, A machine learning-based framework for diagnosis of covid-19 from chest x-ray images, Interdisciplinary Sciences: Computational Life Sciences 13 (2021) 103–117.

[14] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, A. Mohammadi, Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images, Pattern Recognition Letters 138 (2020) 638–643.

[15] W. Zhang, T. Zhou, Q. Lu, X. Wang, C. Zhu, H. Sun, Z. Wang, S. K. Lo, F.-Y. Wang, Dynamic-fusion-based federated learning for covid-19 detection, IEEE Internet of Things Journal 8 (2021) 15884–15891.

[16] B. Pfitzner, N. Steckhan, B. Arnrich, Federated learning in a medical context: A systematic literature review, ACM Transactions on Internet Technology (TOIT) 21 (2021) 1–31.

[17] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[18] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification, Neurocomputing 321 (2018) 321–331.

[19] Zeshan, F. Hussain, D. Gimenez, D. Yi, Rubin, Differential data augmentation techniques for medical imaging classification tasks, PubMed 2017 (2018) 979–984.