An Improvement for Medical Image Analysis Using Data Enhancement Techniques in Deep Learning

Abdulaziz Namozov
Department of Computer Engineering
Gachon University
Seognam-si, Korea Republic
anamozov@gmail.com

Young Im Cho
Department of Computer Engineering
Gachon University
Seognam-si, Korea Republic
yicho@gachon.ac.kr

***Abstract*— A huge number of applications and algorithms have been suggested to analyze medical images. Recent developments in deep learning especially, deep convolutional neural networks (CNN), improved the performance of medical image classification methods. However, training a deep CNN from scratch with medical images is complicated task as it requires a large amount of labelled data. In this paper, we show the role of using different data augmentation techniques to solve such problems. Firstly, we created a deep CNN model with twelve layers for image classification. Then we trained this network with original computed tomography scan (CT) image dataset and some new datasets which are created by generating new images using our original image data. By comparing the classification results of our network on different datasets, we show that using data augmentation techniques can help to improve the medical image classification results and to boost up the network performance**

Keywords—deep learning, medical image analysis, data augmentation, computed tomography images

I. Introduction

Image classification problem is one of the famous tasks in computer vision. Recently, hundreds of research have been done to achieve high performance in image recognition, detection and classification problems. Development of different deep learning algorithms, especially Convolutional Neural Networks brought huge improvements in classification results. For instance, multi-column deep neural network algorithm proposed by Ciregan *et al.* [6] achieved very good performance on MNIST [11], Latin letters [12], Chinese characters [13]. These datasets consist of simple and augmented with label-preserving transformations.

However, achieving such high results in natural image recognition problem is straightforward. To learn thousands of types of objects from millions of images requires very powerful and robust deep learning model with high learning capacity. The first such CNN model was proposed by Alex *et al.* [1] which achieved top-1 and top-5 error rates of 37.5 % and 17 % in ImageNet LSVRC-2010. Afterward, other deep CNN models showed even more excellent classification and detection accuracy in natural image datasets [2] [3] [4] [5].

All of these methods, mentioned above require a huge amount of image data to learn, in order to achieve successful performance. On the other hand, there are some specific tasks such as medical image anaylsis where gathering a large number of images is not possible. Therefore methods based on deep convolutional neural networks results can suffer different problems with small datasets where they cannot generalize well on validation or test sets. The problem that causes bad generalization called over-fitting.

Different approaches have been suggested to reduce over-fitting. Adding regularization term, or using dropout techniques are most common ways of preventing the over-fitting problem. Another popular technique is batch normalization [7]. This method can help not only for learning training data successfully, it can also decrease computational expenses. Another powerful method that we analyze in this paper is increasing training data using different data augmentation techniques. We show that by using different kinds of data augmentation techniques, we can improve the classification accuracy of deep CNN model for CT image classification.

The rest of the paper is organized as follows: Section II gives a brief review of different data augmentation techniques. The architecture of our deep CNN network for medical CT image classification and data enhancement techniques that we used in our problem are explained in Section III. Experimental results are demonstrated in Section IV. And we conclude our work in Section V

II. Related Work

The area of data augmentation is not new and many types of data enhancement methods have been applied to specific problems. The main techniques fall under the category of data warping, which is an approach which seeks to directly augment the input data to the model in data space. The idea can be traced back to augmentation performed on the MNIST set in [8].

A very generic and accepted current practice for augmenting image data is to perform geometric and color augmentations, such as reflecting the image, cropping and translating the image, and changing the color palette of the image. All of the transformations are the affine transformation of the original image that takes the form.

This idea has been carried further in [9], where an error rate of 0:35% was achieved by generating new training samples using data augmentation techniques at each layer of a deep network. Specifically, digit data was augmented with elastic deformations, in addition to the typical affine transformation. Furthermore, data augmentation has found applicability in areas outside simply creating more data.

Generative Adversarial Nets (GANs) has been a powerful technique to perform unsupervised generation of new images for training. They have also proven extremely effective in many data generation tasks, such as novel paragraph generation [10]. By using a min-max strategy, one neural net successively generates better counterfeit samples from the original data distribution in order to fool the other net. The other net is then trained to better distinguish the counterfeits. GANs have been used for style transfer such as transferring images in one setting to another setting (CycleGAN) [14].

III. Proposed Approach

A. Data Augmentation

We show the effectiveness of augmentation techniques to solve the over-fitting problem, by training a deep CNN model in the CT image classification problem. We created a small dataset of CT images of brain. Originally, our dataset consists 500 images (250 positive and negative) as shown in **Table 1.** And we increased the number of only *train images* using data augmentation techniques.

TABLE I. Different image datasets

Datasets	Number of positive images	Number of negative images	Total number of images	Purposes for
Set1	250	250	500	Training
Set2	1000	1000	2000	Training
Set3	50	50	100	Testing

Augmenting image data using traditional transformation techniques includes combinations of affine transformations or flipping the image horizontally or vertically, or rotating by 90 degrees. In this work, for any input image we created other new 3 images using horizontal and vertical flip and rotation by 90 degrees **(Figure 1)**. For original dataset (*Set1*) of *N* size, we generate a dataset of *4N* size. We call this dataset *Set2* as shown in **Table 1.**

B. Network Architecture

As we just try to show the effectiveness of using data augmentation for reducing the over-fitting problem, we created deep CNN architecture inspired by VGG-Net [2] which consist of 12 layers as shown in **Figure 2.**

We have created a deep CNN for medical CT image classification using some of the ideas fundamental to VGG-Net [2], which showed excellent results in the classification task

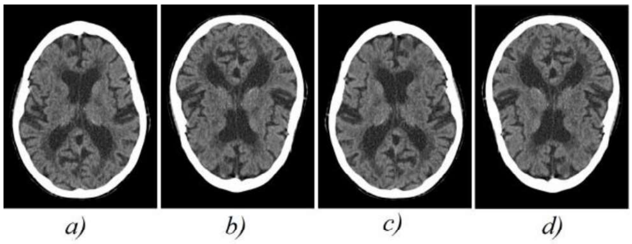

Fig. 1. An example of data augmentation. a) Original image in training set; b) A new image generated by horizontal flip; c) A new image by generated by vertical flip; d) A new image rotated by 90 degrees.

of the 2014 ImageNet Challenge. Considering the fact that that we only solve a two-class problem, we create a network that is not deep as the original VGG-net; it consists of 12 layers, as shown in Fig. 2. The network has six convolutional and three pooling layers for feature extraction, and three fully-connected layers, in the end, for classification.

1) Convolutional layers

We use small-size 3×3 convolutional filters for all convolutional layers. Each of first two convolutional layers has 32 feature maps, and after each pooling layer, we double the number of filters. Therefore, the third and fourth convolutional layers include 64 feature maps, whereas there are 128 feature maps in each last convolutional layer.

2) Pooling layers

Feature maps in convolutional layers are used to extract features. The role of pooling layers in CNNs is to attain special invariance by reducing the resolution of those feature maps. This can help to decrease the computational time, as well as to address the risk of the over-fitting problem. Traditionally, pooling layers are connected to convolutional layers. Average pooling or max pooling is most commonly used in deep learning models. The former uses the average activation value over a pooling region, whereas the latter selects the maximum activation value. The max pooling technique is used in our network. We apply the same method for all three pooling layers over a 2×2 pixel window with stride 2.

3) Fully-connected layers

After convolutional and pooling layers, we add three fully-connected layers to perform the classification. Generally, fully-connected layers include most of the learnable parameters. In our case, the first two fully-connected layers have 2048 channels each, and the third layer (which is the output layer of the network) contains three neurons and performs 2-way positive and negative medical image classification. To prevent over-fitting, we use the dropout technique in the first two fully-connected layers. We use the softmax function, which computes the probability for each class in the third fully-connected layer.

C. Network Training

We trained the neural network model with 2 different datasets: original dataset *Set1* with no augmentation, *Set2* dataset, which was generated using traditional augmentation techniques, And *Set3* was used for testing for all cases mentioned above.

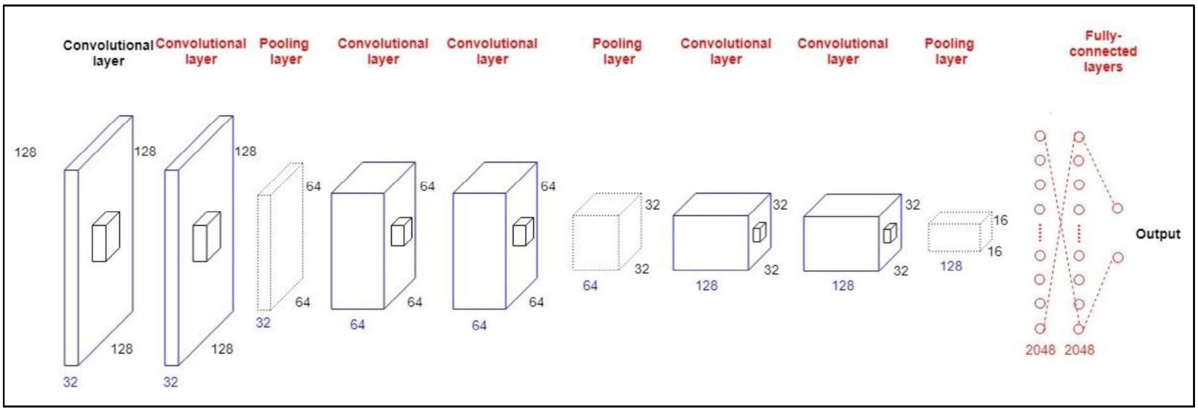

Figure 1. The architecture of the deep CNN for medical image classification. The network consists of twelve layers: six convolutional layers (cuboids with blue lines), three max-pooling layers (cuboids with dotted lines) and three fully connected layers at the end. Blue numbers under cuboids denote the number of filters, black numbers denote the widths and heights of the feature maps and red numbers denote the number of neurons in the fully-connected layers.

IV. Experiments and Results

We used Keras [15] to build and train our CNN model. All experiments were carried out in the Windows 10 operating system running on a PC with *Intel (R) core i7-770HQ* CPU and Nvidia GeForce GTX 1050Ti.

TABLE II. RESULTS ON MEDICAL IMAGE CLASSIFICATION

Datasets	Train accuracy (%)	Test accuracy (%)
Set1	85.725	72.611
Set2	93.157	92.25

Experimental results in Table 2 show that when the model trained with only original images, train and test accuracy after 100 epochs were the worst with only almost 86 and 73 % respectively. It means, the network is suffering from the over-fitting problem, where the difference between train and test accuracy was enormous. However, applying augmentation techniques reduced the over-fitting and considerably improved both train and test accuracy. The dataset with traditional augmentation techniques achieved 93.15% train and 92.25% test and test accuracy.

V. Conclusion

In this paper, we analyze the usefulness of data augmentation techniques to improve the performance of the deep CNNs for medical image analysis. Experiments in CT medical image classification task show that when we improve the number of train images using some image generation methods, the network model can generalize better on validation data. Our experiments can be base to achieve better generalization in medical image recognition and classification problems, in particular when training samples are insufficient and imbalanced.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2018-2017-0-01630) supervised by the IITP(Institute for Information & communications Technology Promotion)

References

[1] 1. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105. 2012.

[2] 2. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[3] 3. Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." Cvpr, 2015.

[4] 4. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." In *European conference on computer vision*, pp. 818-833. Springer, Cham, 2014.

[5] 5. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

[6] 6. Ciregan, Dan, Ueli Meier, and Jürgen Schmidhuber. "Multi-column deep neural networks for image classification." *Computer vision and pattern recognition (CVPR), 2012 IEEE conference on*. IEEE, 2012.

[7] 7. Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *arXiv preprint arXiv:1502.03167*. 2015.

[8] 8. Baird, Henry S. "Document image defect models." *Structured Document Image Analysis*. Springer, Berlin. 546-556, 1992

[9] 9. D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber. Deep big simple neural nets excel on handwritten digit recognition. CoRR, abs/1003.0358, 2010.

[10] 10. Liang, Xiaodan, Zhiting Hu, Hao Zhang, Chuang Gan, and Eric P. Xing. "Recurrent topic-transition GAN for visual paragraph generation." *CoRR, abs/1703.07022* 2, 2017.

[11] 11. LeCun, Yann, Corinna Cortes, and C. J. Burges. "MNIST handwritten digit database." *AT&T Labs [Online]. Available: http://yann.lecun. com/exdb/mnist* 2, 2010.

[12] 12. P. J. Grother. NIST special database 19 - Handprinted forms and characters database. Technical report, National Institute of Standards and Technology (NIST), 1995

[13] 13. Liu, Cheng-Lin, Fei Yin, Da-Han Wang, and Qiu-Feng Wang. "Chinese handwriting recognition contest 2010." In *Pattern Recognition (CCPR), 2010 Chinese Conference on*, pp. 1-5. IEEE, 2010.

[14] 14. Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *arXiv preprint arXiv:1703.10593* 2017.

[15] https://keras.io/