Proceedings of the SICE Annual Conference 2016
Tsukuba, Japan, September 20-23, 2016

Mass Detection Using Deep Convolutional Neural Network for Mammographic Computer-Aided Diagnosis

Shintaro Suzuki[1†], Xiaoyong Zhang[1], Noriyasu Homma[2], Kei Ichiji[2], Norihiro Sugita[1], Yusuke Kawasumi[2], Tadashi Ishibashi[2], and Makoto Yoshizawa[3]

[1]Graduate School of Engineering, Tohoku University, Sendai, Japan
(Tel: +81-22-795-7130; E-mail: shintaro.suzuki.r4@dc.tohoku.ac.jp)
[2]Tohoku University Graduate School of Medicine, Sendai, Japan
[3]Cyberscience Center, Tohoku University, Sendai, Japan

Abstract: In recent years, a deep convolutional neural network (DCNN) has attracted great attention due to its outstanding performance in recognition of natural images. However, the DCNN performance for medical image recognition is still uncertain because collecting a large amount of training data is difficult. To solve the problem of the DCNN, we adopt a transfer learning strategy, and demonstrate feasibilities of the DCNN and of the transfer learning strategy for mass detection in mammographic images. We adopt a DCNN architecture that consists of 8 layers with weight, including 5 convolutional layers, and 3 fully-connected layers in this study. We first train the DCNN using about 1.2 million natural images for classification of 1,000 classes. Then, we modify the last fully-connected layer of the DCNN and subsequently train the DCNN using 1,656 regions of interest in mammographic image for two classes classification: mass and normal. The detection test is conducted on 198 mammographic images including 99 mass images and 99 normal images. The experimental results showed that the sensitivity of the mass detection was $89.9\,\%$ and the false positive was $19.2\,\%$. These results demonstrated that the DCNN trained by transfer learning strategy has a potential to be a key system for mammographic mass detection computer-aided diagnosis (CAD). In addition, to the best of our knowledge, our study is the first demonstration of the DCNN for mammographic CAD application.

Keywords: Neural Networks, Signal and/or Image Processing, Medical and Welfare Systems, Deep Learning, DCNN, Transfer Learning, mammogram, Computer-Aided Diagnosis/Detection

1. INTRODUCTION

For years, cancer has been one of the biggest threats to human life. In Japan, breast cancer is the most frequent cancer among women in recent years, and the number of patients is continue to increase [1]. In general, early detection is very important for cancer treatment and accurate detection of breast cancer in its early stages can give a better chance of full recovery. For this purpose, it is recommended for women to have breast cancer examination using mammography [2]. Mammography is a specific type of X-ray imaging modality that is effective to detect early breast cancer before it becomes clinically palpable [3]. Today, the mammography has widely been accepted for routine breast cancer screening. However, due to the increase of the number of examinee, reading a lot of mammograms became burden for doctors, and it might cause misdiagnosis that leads to unnecessary biopsies or mastectomies.

For reducing the work burden of doctors and improving their detection accuracy, computer-aided diagnosis (CAD) systems that use computer technologies to detect abnormalities in mammograms, have been developed [4]. The CAD systems are almost focus on detecting 3 typical abnormalities in mammograms: microcalcification cluster (MCC), mass, and architectural distortion (AD). In this paper, our focus is on detection of mass because the mass is one of the most important mammographic signs of malignancy [5]. On mammograms, mass usually has a circular shape and it has tendency to be denser in the middle compared to its edge.

Most of the previous mass detection methods [6–9] are based on the following three phases: 1) mass candidate or region of interest (ROI) detection using image processing techniques; 2) extracting the feature vector of the ROI based on special knowledge; 3) classifying the ROI into normal or mass classes. However, it is difficult to design the features that can sufficiently represent the characteristics of masses to distinguish the masses from the others.

In recent years, deep learning has attracted great attention in artificial intelligence (AI) due to its successes in various research fields, such as pattern recognition, computer vision and big-data analysis. As one of the most successful techniques in deep learning, DCNN achieved outstanding performance in recognition of natural images [10]. The DCNN can obtain feature representation and classification rules from training data automatically, without hand-designed feature extraction. However, the DCNN performance for mammographic image recognition is still uncertain because collecting a sufficiently large amount of training data is difficult. In addition, due to the difference of patterns between the natural image and the mammogram image, the ability of the DCNN in obtaining the feature representation is also uncertain.

In this paper, we present a transfer learning strategy to solve the above training problem of the DCNN for a mammographic CAD application. In this strategy, a huge

† Shintaro Suzuki is the presenter of this paper.

978-4-907764-50-0 PR0001/16 ¥400 ©2016 SICE

amount of natural images are firstly used for pre-training, and then a relatively small number of mammographic images are used for transfer learning. Experimented results showed that the DCNN after the transfer learning is capable of perceiving the difference between the mass and normal region.

2. PROPOSED METHOD

In this study, we adopt a previous DCNN architecture which is called *AlexNet* [10] for our application.

2.1. Architecture of the DCNN

As shown in Fig.1, the DCNN contains 8 layers with weights: the first 5 layers are convolution layers and the remaining 3 layers are fully-connected layers. The input of the DCNN is the intensity volume (R,G,B) of an image. The output of the DCNN produces a distribution of predicted probability over the 1000 classes for *ImageNet* classification [11].

Here, we briefly describe each layer of the DCNN, and the details be found in [10].

The layers of *conv1-5* in Fig. 1 are the convolution layers. Each neuron in the convolution layers computes a dot product between their weights and the local region that is connected to in the input volume. The weights consist of a set of learnable filters. Every filter is small spatially (along width and height), but extends through the full depth of the input volume. We slide each filter across the width and height of the input volume, producing a 2-dimensional activation map of that filter. All neurons in a single 2-dimensional slice of depth are using the same weight vector, hence the output is computed as a convolution of the neuron's weight with the input volume. Adjusting the parameters (weights and biases of neurons) by training, the DCNN will learn filters that activate when they see some specific type of feature at some spatial position in the input.

The layers of *pool1*, *pool2* and *pool5* in Fig. 1 are the pooling layers. The pooling layer performs a downsampling operation along the spatial dimensions to reduce the amount of computation and improve the robustness to position variation. We adopted the pooling layers with filters of size 3×3 pixels. Each neuron in these layers outputs the maximum value of 9 input in each local region. And we shift the filters 2 pixels at a time to reduce the width and height of input volume to half. The depth dimension remains unchanged.

The layers of *norm1* and *norm2* in Fig. 1 are the normalization layers. We used a local response normalization (LRN) layer [10] in this study. It performs a kind of "lateral inhibition" that observed in the brain, by normalizing over multiple channels at the same position in the input. Its function is to encourage competition for too large activation.

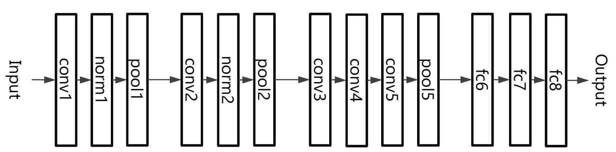


Fig. 1: The DCNN architecture.

Table 1: Confusion matrix for mass classification after 30 training epochs. The DCNN achieved 89.9 % sensitivity and 19.2 % false positive in mass detection.

		Actual class		
		Mass	Normal	Total
Predicted class	Mass	89	19	108
	Normal	10	80	90
	Total	99	99	198

The layers of *fc6*, *fc7* and *fc8* in Fig. 1 are the fully-connected layers. Neurons in the fully-connected layer have full connections to all neurons in the previous layer, as seen in ordinary feedforward neural networks. The layer of *fc8* consists of 1000 neurons, and each neuron outputs the class scores that represent the predicted probability that the input is belongs to each class of *ImageNet* classification [11].

2.2. Transfer Learning

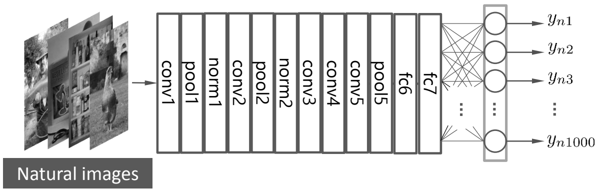
It is known that the extracted features of earlier layers of the DCNN contain more generic features (e.g. edge detectors or color blob detectors) that should be useful to many tasks, but in later layers, generic features are combined and become more specific to the details of the classes contained in the training dataset. For this reason, it is effective to pre-train the DCNN by using very large dataset, and then use the DCNN as a feature extractor for other task of which the amount of training data is too small [12].

In this study, we used a fine-tuning strategy to transfer learned recognition capabilities from general task (classification of natural images) to the specific task of recognition of masses on mammograms. The fine-tuning strategy is a type of transfer learning scenarios, and it is different to other scenarios in continuing to train the pre-trained DCNN using the dataset of the task of interest.

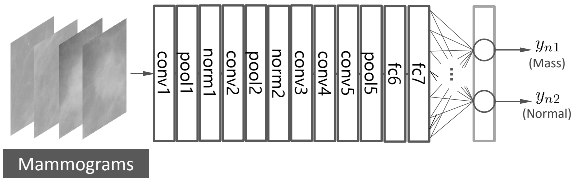
As shown in Fig. 2(a), the DCNN is firstly pre-trained by using *ImageNet* dataset [11], which contains 1.2 million natural images for classification of 1,000 classes. Then, the last fully-connected layer is replaced by a new layer for classification of 2 classes: mass and normal. The modified DCNN is subsequently retrained by a small-scale mammographic image dataset as shown in Fig. 2(b).

3. EXPERIMENTAL RESULTS

The transfer training and test experiments are conducted on the Digital Database for Screening Mammography (DDSM) [13]. The training and test image data



(a) Pre-training by using natural images.



(b) Fine-tuning by using mammograms.

Fig. 2: The fine-tuning transfer learning strategy. (1) The DCNN is pre-trained by using a large amount of images, then we modified the last layer into two neurons, corresponding mass and normal classes respectively, and subsequently (2) we retrained the DCNN on a small-scale mammographic image dataset.

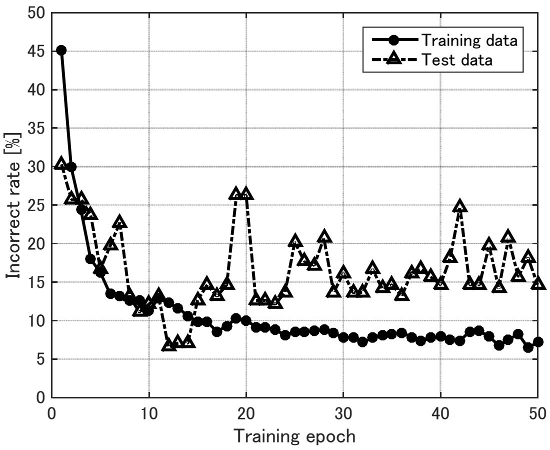


Fig. 3: Changes of classification error rate. The incorrect rate for test data decreased and finally reached about 15 %.

consist of two groups of ROI images cropped from the original mammograms. The first group ROI contains the mass area and the second one is the normal area. Both ROIs are size of 454×454 pixels. In the transfer training phase, 1,656 ROI images including 786 mass images and 870 normal images are used to retrain the DCNN. The test experiment was conducted on 198 mammographic ROI images including 99 mass images and 99 normal images.

Fig. 3 shows that the incorrect rate of mass detection for training data (solid line) gradually decreased with the training epoch, and the incorrect rate for test data (broken line) also decreased and reached about 15% after 30 training epochs. As summarized in Table 1, the DCNN finally achieved 89.9% sensitivity and 19.2% false positive in mass detection after 30 training epochs. These results demonstrated that the DCNN is capable of classifying the mass from normal ROI in some extent.

Fig. 4(a) shows the top 28 images which are classified into mass in the test experiment. The number under each image is the class score that corresponding to a predicted probability that the input is a mass image. On the other hand, Fig 4(b) shows the images which are classified into normal.

Fig. 5(a) shows the normal images that were misclassified into mass (false positive). In these images, we can find a high brightness pattern of healthy tissue, which is similar to the pattern of mass. Fig. 5(b) shows the overlooked mass images. Some of these masses are indistinct and others are too large and protruded from the ROI.

4. CONCLUSION

In this study, we presented a transfer learning of the DCNN for mass detection in mammogram images. The experimental results demonstrated that the DCNN and transfer learning strategy have a promise potential for mammographic CAD system even the number of training data is limited. To the best of our knowledge, this study is the first demonstration of DCNNs for detecting the masses in mammographic images.

5. ACKNOWLEDGEMENT

This work was partially supported by JSPS KAKENHI Grant No.15K20852 and No.26540112.

REFERENCES

[1] A. Matsuda, T. Matsuda, A. Shibata, K. Katanoda, T. Sobue, H. Nishimoto and The Japan Cancer Surveillance Research Group, “Cancer Incidence and Incidence Rates in Japan in 2008: A Study of 25 Population-based Cancer Registries for the Monitoring of Cancer Incidence in Japan (MCIJ) Project,” *Japanese Journal of Clinical Oncology*, **44**-4, pp. 388-396, 2013.

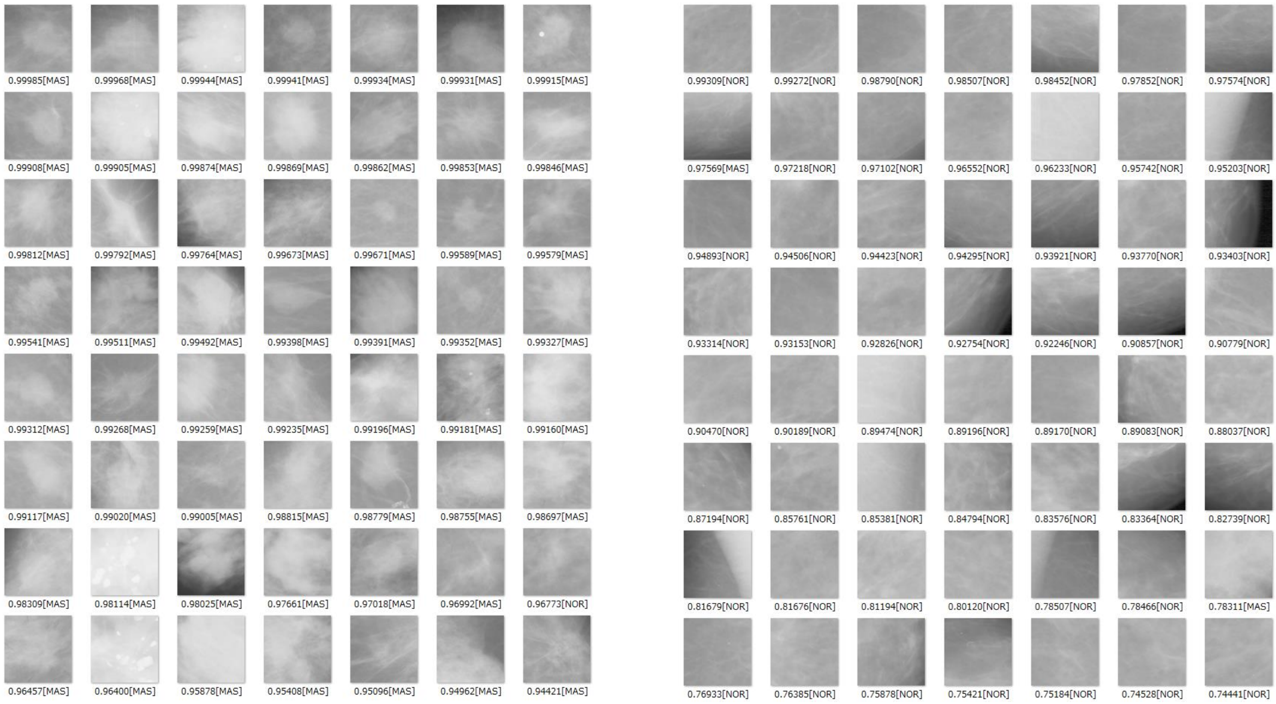
[2] N. J. Wald, J. Chamberlain, and A. Hackshow, “Consensus statement : Report of the European Society for Mastology Breast Cancer Screening Evaluation Committee,” *Breast*, **2**, pp. 209-216, 1993.

[3] K. H. Ng and M. Muttarak, “Advances in Mammography Have Improves Early Detection of Breast Cancer,” *JOURNAL-HONG KONG COLLEGE OF RADIOLOGISTS*, **6**, pp. 126-131, 2003.

[4] J. Tang, R. M. Rangayyan, J. Xu, I. E. Naqa, and Y. Yang, “Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography: Recent Advances,” *IEEE Trans. Inf. Technol. Biomed*, **13**-2, pp. 236-251, 2009.

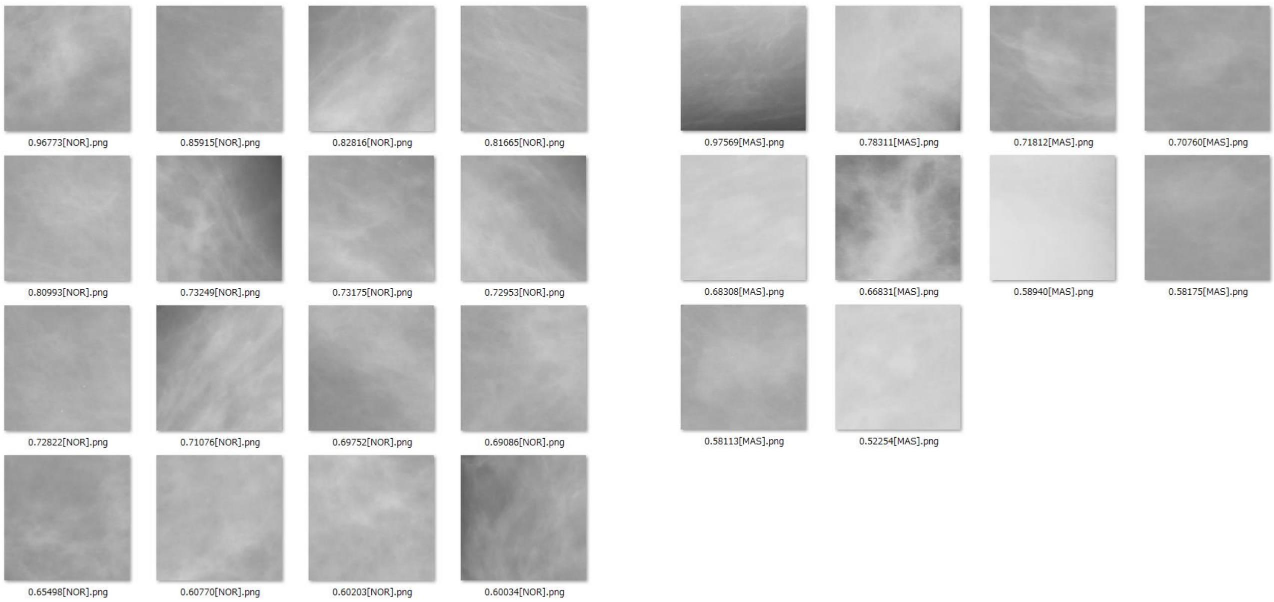
[5] B. Sahiner, N. Petrick, H. P. Chan, L. M. Hadjiiski, C. Paramagui, M. A. Helvie, and M. N. Gurcan, “Computer-Aided Characterization of Mammographic Masses: Accuracy of Mass Segmentation and Its Effects on Characterization,” *IEEE Trans. Med. Img*, **20**-12, pp. 1275-1284, 2001.

[6] H. P. Chan, D. Wei, M. A. Helvie, B. Sahiner, D. D. Adler, M. M. Goodsit, and N. Petrick, “Computer-



(a) (b)

Fig. 4: The classified images with a predicted probability. (a) The mass classification with probability ranking. (b) The normal classification with probability ranking.



(a) (b)

Fig. 5: The misclassfied images. (a) False positive in mass classification. (b) False negative in mass classification.

aided classification of mammographicmasses and normal tissue: Linear discriminant analysis in texture feature space," *Phys. Med. Biol.*, **40**-5, pp. 857-876, 1995.

[7] B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolutional neural network classifier with spatial domain and texture images," *IEEE Trans. Med. Imag.*, **15**-5, pp. 598-610, 1996.

[8] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels, Gradient and texture analysis for the classification of mammographic masses, *IEEE Trans. Med. Imag.*, **19**-10, pp. 1032-1043, 2000.

[9] J. Wei, B. Sahiner, L. Hadjiiski, H. Chan, N. Petrick, M. Helvie, M. Roubidoux, J. Ge, and C. Zhou,

Computer aided detection of breast masses on full field digital mammograms, *Med. Phys.*, **32**-9, pp. 2827-2837, 2005.

[10] A. Krizhevski, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," In *Proc. NIPS*, 2012.

[11] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," In *CVPR09*, pp. 248-255, 2009.

[12] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrel, "DeCAF: A deep convolutionalactivation feature for generic visual recognition," *JMLR W&CP*, **32**-1, pp. 647-655, 2014.

[13] M. Heath, K. Bowyer, D. Kopans, R. Moore and W. P. Kegelmeyer, "The Digital Database for Screening Mammography," In *Proceedings of the Fifth International Workshop on Digital Mammography*, M.J. Yaffe, ed., pp. 212-218, Medical Physics Publishing, 2001.