

Deep Learning for Cancer Screening in Medical Imaging

Jihoon Jeong^{1,2,3}

¹Advisor, Lunit Inc., Seoul, Korea

²Partner, Digital Healthcare Partners, Seoul, Korea

³Senior Teaching Fellow, Department of Media and Communications, Kyung Hee Cyber University, Seoul, Korea

In recent years, deep learning has been used in many researches in cancer screening based on medical imaging. Among cancer screening using optical imaging, melanoma detection is the biggest concern. Stanford University researchers used CNNs (convolutional neural networks) to classify skin lesions comparing with 21 dermatologists for 2 tasks. CNN performed better than all the dermatologists' tasks. Finding pulmonary nodules on chest X-ray has the longest history in cancer screening using medical imaging and neural network technology began to be applied before the deep learning technology matured as it is now. But, the applications were mainly focused on screening in CT images. There is relatively few research on pulmonary nodule detection using deep learning in chest X-rays. For breast cancer screening in mammography, adoption of neural network technologies has already begun early. Many studies have shown that tumor detection using CNNs is useful in breast cancer screening. Most of the results are from mammography, but studies using tomosynthesis, ultrasound, and MRI have also been published. Although imaging modality and target cancer are different, we can see that there are similar kinds of future challenges. First, it is not easy to acquire a large amount of medical image data required for deep learning. Second, it is difficult to learn if there are many medical image data but they are not properly labeled. Finally, there is a need for technologies that can use different imaging modalities at the same time, link with electronic health records, and use genetic information for more comprehensive screening.

Key words: Deep Learning; Convolutional Neural Networks; CNNs; Computer Aided Diagnosis; CAD; Skin Cancer; Mammography; Chest X-ray

INTRODUCTION

Cancer screening in medical imaging is one of the most important areas in computerized medical software. Especially, attempts to automate the early diagnosis of cancer using computer aided detection (CAD) algorithm on chest X-ray and mammography images were the most important research topic in the field of radiology [1]. However, the results of the clinical effects of CAD are still controversial. Even there was a research about screening performance of CAD reporting that sensitivity was significantly decreased for mammograms interpreted with

vs without CAD in the subset of radiologists who interpreted both with and without CAD (odds ratio, 0.53; 95% CI, 0.29-0.97) [2]. But, deep learning technology, which has recently been greatly developed, is raising expectations for the possibility of computer software related to cancer screening again.

Deep learning is a kind of neural network. The neural network consists of an input layer, a hidden layer, and an output layer. Deep learning is a neural network with a large number of hidden layers. Over the past few years, deep learning has achieved tremendous performance improvements, especially in image classification [3] and speech recognition [4]. In recent

Corresponding Author: Jihoon Jeong

Advisor, Lunit Inc., 6th Floor, 175 Yeoksam-ro, Gangnam-gu, Seoul, Korea
Tel: +82-10-2512-2540
E-mail: jjjeong@lunit.io

Received 25 July 2017

Revised 18 Sep 2017

Accepted 2 Nov 2017

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

years, deep learning has been used in a wide variety of areas. In particular, powerful image recognition and classification capabilities are being continuously adopted in high-tech devices such as autonomous vehicles and drones. The medical sector is no exception. Google has reported good results using deep learning for diagnosis of diabetic retinopathy [5], and Mount Sinai Hospital has introduced a system called “Deep Patient” that applies this technology throughout the electronic health records (EHRs) system, reporting improved accuracy in predicting the prognosis of severe diabetes, schizophrenia, and various cancers [6].

In this review, I would like to describe what the deep learning technology has done to cancer screening, what features of applied technology are, and what challenges to overcome in the future, focusing on the medical imaging field. As another review of this issue addresses CT, MRI and pathology applications, this review focuses on key cases that have been applied to optical imaging, chest X-ray, and mammography.

CONVOLUTIONAL NEURAL NETWORKS

Of the deep learning technologies, convolutional neural networks (CNNs) are particularly popular in the field of medical imaging. Most of the deep learning technologies related to cancer screening are also based on CNNs. CNN has been growing since the late 1970s and has been applied to medical imaging analysis since 1995 [7]. However, CNNs were able to be applied in the real world since Yann LeCun's LeNet was applied to hand-written digit recognition [8]. But, other than that, CNNs have become a technology with success not seen by many until 2012. The event that broke the downturn of CNNs was the December 2012 ImageNet Challenge. Krizhevsky et al, using CNNs called AlexNet, became the winners of the competition in a remarkable gap [9]. AlexNet contains eight layers with weights. The first five are convolutional and the remaining three are fully connected layers. Since the success of AlexNet, numerous CNNs architectures have emerged and continue to boost performance. Finally, the 2015 ImageNet Challenge comes with ResNet, a CNN that achieves an error rate of 3.57%, which is below the human average of 5% [10]. ResNet makes use of the residual function to enable deep learning of up to 152 layers, and since then there have been many similar CNNs with very good performances. From this point on, deep learning research using CNNs has started to

increase rapidly in the field of medical imaging.

CANCER SCREENING IN OPTICAL IMAGING

Among cancer screening using optical imaging, melanoma detection is the biggest concern. According to The American Cancer Society, 87,110 new melanomas are diagnosed in the United States in 2017 and 9,730 are estimated to die. In addition, the 5-year survival rate of melanoma diagnosed as Stage IA and IB was 97% and 92%, respectively [11]. This means that although the problem is serious, early diagnosis can save many people. Therefore, early diagnosis of melanomas using optical imaging is very important. One of the most remarkable achievements of cancer screening in medical imaging using deep learning has come from melanoma detection. Stanford University researchers used CNNs to classify skin lesions. They used 129,450 clinical images to train classifying 2,032 different diseases and compared the test results with the results read by 21 board-certified dermatologists. The tasks tested in this study were to differentiate malignant carcinomas versus benign seborrheic keratoses and malignant melanomas versus benign nevi. In this experiment, CNN performed better than all dermatologists in all tasks [12]. This shows the potential to distinguish malignant skin cancer early from a photograph taken through a smart phone, and it has shown that cancer screening using optical imaging can be an inexpensive and efficient medical service.

In addition to dermatological diseases, cancer screening using optical imaging is expected to be useful for cervical cancer diagnosis. In 2016, several conference papers published improved results on cervical dysplasia diagnosis and cervical cancer classification [13]. Because cervical cancer ranks as the second most common type of cancer in women aged 15 to 44 years worldwide, these results have very important implications for women's health. To facilitate this research, Intel and MobileODT are collaborating on a cervical cancer screening challenge through Kaggle, a data platform. This challenge began on March 16, 2017 and ended on June 21, 2017, with 848 teams from 22 countries around the world participating in a hot competition [14]. Because the result of such a challenge is made with data only, it is necessary to carry out rigorous clinical experiments in order to become a practical applicable technology. However, the pace of development of this technology is expected to be very fast, because already top teams are releasing and discussing their technology.

CANCER SCREENING IN CHEST X-RAY

Finding pulmonary nodules on chest X-ray has the longest history in cancer screening using medical imaging. Neural network technology began to be applied before deep learning technology matured as it is seen now [7]. The reason for this early adoption of the new technology is that early diagnosis of lung cancer using chest X-ray has such an important clinical significance. According to the 2017 statistics of the American Cancer Society, there are about 222,500 cases estimated to be newly diagnosed with 155,870 deaths in 2017. This is the number one cause of cancer deaths regardless of gender, and it reaches one-quarter of all cancer deaths [15].

However, studies using deep learning techniques, including CNNs, for pulmonary nodule detection have begun to emerge after 2015. But, the applications were mainly focused on screening in CT images. There is relatively few research on pulmonary nodule detection using deep learning in chest X-rays. Bar et al used CNNs to identify different types of pathologies in chest X-ray images. They tested a dataset of 93 images trained with ImageNet, a well-known large scale nonmedical image database and try various combinations of features extracted from the CNN and a set of low-level features. They got an area under the curve (AUC) of 0.79 for classification between healthy and abnormal chest X-rays, where all pathologies are combined into one large class [16]. Bobadilla et al tried a family of CNNs known as DeepCNets composed of alternating convolutional and max-pooling layers with a linearly increasing number of filters. They trained and tested the Japanese Society of Radiological Technology (JSRT) database [17] composed of 154 nodules and 93 non-nodule chest radiographs. Their results show that CNNs can operate effectively on lung nodule classification through data augmentation and dropout regularization and achieve comparable results with state-of-the-art performance [18].

There are far fewer deep learning studies using chest X-rays compared to CT because there are not enough public datasets to accelerate this research. Conversely, in the case of CT screening, challenges such as LUNA16 or Kaggle Data Science Bowl 2017, which provide rich public datasets and large prize pools, are nurturing deep learning research. However, it is clear that the chest X-ray is the simplest and cheapest modality. Therefore, deep learning studies for lung cancer screening using chest X-rays will still be very important.

CANCER SCREENING IN MAMMOGRAPHY

The application of neural network technology for breast cancer screening in mammography has already begun in 1996 [19]. Also, in 2012, Jamieson et al. attempted mass classification of mammography and ultrasound using 4-layer AND [20]. In this way, compared with other cancers, breast cancer screening was applied relatively early. This is probably due to the importance of breast cancer as the number one female cancer mortality rate and commonly used cancer screening imaging modality such as mammography. However, deep learning has been widely applied in this field since 2015. Many studies have shown that tumor detection using CNNs is useful in breast cancer screening. Most of the results are from mammography [21,22,23,24], but studies using tomosynthesis [25,26], ultrasound [27], and MRI [28] have also been published.

Although the results so far have shown promise, there are still no results showing state-of-the-art results that are well beyond radiologists. The main reason is that, like other medical imaging deep learning studies, it is difficult to acquire a large amount of data. Therefore, various studies are under way to overcome these limitations.

FUTURE CHALLENGE AND PERSPECTIVES

So far, I have reviewed the current state of deep learning technology used in cancer screening focusing on optical imaging, chest X-ray, and mammography. Although imaging modality and target cancer are different, we can see that there are similar kinds of future challenges.

First, it is not easy to acquire a large amount of medical image data required for deep learning. Therefore, most of the research results utilize the deep neural network used for other image recognition or additional training after transfer learning from other networks. To solve this problem, it is important to have a large number of public medical image datasets like ImageNet. Challenges such as Intel&MobileODT Cervical Cancer Screening and LUNA that target public competition through Kaggle and lung nodule detection are some of them, but more fundamental efforts to solve the problems should be done. Recently, Stanford University has announced plans that will build Medical ImageNet and activate research on medical imaging deep learning based on this plan. This effort should have global consensus and be able to collaborate with multiple medical institutions.

Second, it is difficult to learn if there are many medical image data but they are not properly labeled and annotated. Deep learning technology, which is applied to most medical images to date, is highly dependent on radiologists' labels and annotations. This is a fundamental problem with supervised learning technology. And, it seems that it may not be solved by unsupervised learning. Therefore, deep learning technology, which can provide good performance with a small number of correct answers, is actually very important. Hwang and Kim proposed a technique called weakly supervised learning to overcome this problem. They used class activation maps from CNNs for weakly-supervised learning with good localization and screening performance [22]. We can use traditional fully-supervised learning if we can obtain medical image data with enough labels and annotation, and weakly-supervised learning if we have no annotations. Semi-weakly-supervised learning can be applied in between. There is a difference between these approaches (Fig. 1). Depending on the data acquisition and performance of the training and test, we need to reasonably determine α (percentage of annotated medical images).

Finally, there is a need for technologies that can use different imaging modalities at the same time, link with EHR, and use genetic information for more comprehensive screening. This is a very difficult task, technically and clinically, but this goal must be accomplished through common effort. For this goal, among

the deep learning technologies, CNNs that are mainly used in medical imaging, RNNs (Recurrent Neural Networks) capable of natural language processing, and reinforcement learning that can be continuously learned in various situations will be necessary together.

CONCLUSION

Cancer screening in medical imaging is a field that can achieve many good results with deep learning technology. As briefly reviewed here, deep learning technologies applied to optical imaging, chest X-ray, and mammography have flourished since 2015 and many studies have been published. Reviewed papers are summarized in Table 1. So far, the results are not very satisfactory, but the development of deep learning technology is very fast, the supply of researchable medical image data is getting bigger, and the research funds are getting rich, so the future is bright. In the future, it will be easier and more accurate to diagnose not only medical images but also EHR and genetic information with the help of deep learning technology. The development of deep learning technologies is important for this, but the role of physicians who understand and use these technologies becomes increasingly important. Therefore, I expect more physicians to study these deep learning technologies with interest.

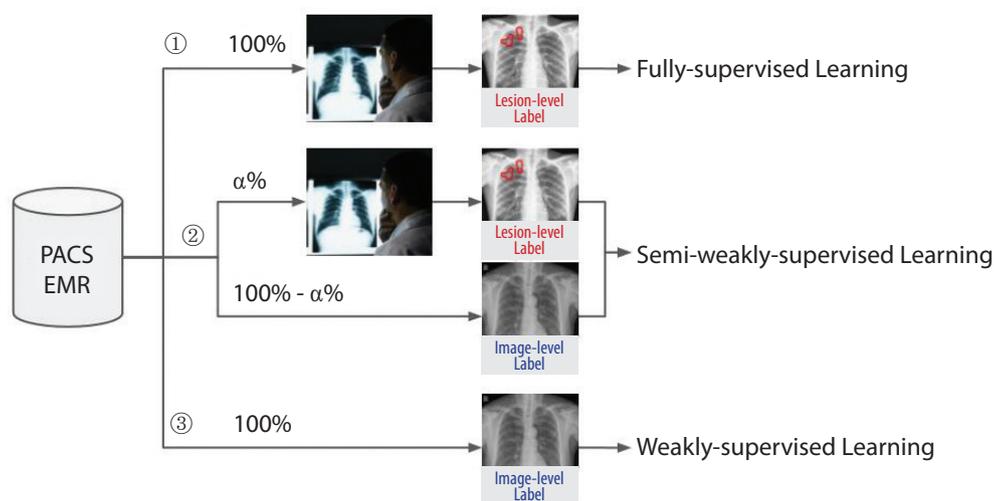


Fig. 1. Three types of learning scenarios. Fully-supervised learning if we can obtain medical image data with full annotations, weakly-supervised learning if we have no annotations. Semi-weakly-supervised learning can be applied in between. We need to determine α (percentage of annotated medical images) reasonably depending on the data acquisition and performance of the training and test.

Table 1. Researches on Cancer Screening using Medical Imaging Deep Learning

Imaging Modality	Reference	Application	Remark
Optical Imaging	Esteva A et al (2017)	Skin Cancer	Distinguish malignant skin cancer early from a photograph taken through a smart phone
	Xu T et al (2016)	Cervical Cancer	Multimodal deep learning application for cervical dysplasia
Chest X-ray	Lo SB et al (1995)	Lung Cancer	Very early application of CNN in medical imaging
	Bar Y et al (2015)	Lung Cancer	Screening for non-medical trainer
	Bobadilla JCM et al (2017)	Lung Cancer	SOTA result on Japanese Society of Radiological Technology (JSRT) databases
Mammography	Sahiner B et al (1996)	Breast Cancer	Earliest application of CNN in mammography
	Jamieson et al (2012)	Breast Cancer	Classification through mammography and ultrasound using 4-layer ADN
	Hwang S et al (2016)	Breast Cancer, Tuberculosis	Self transfer-learning and weakly-supervised learning technique introduced
	Wang J et al (2016)	Breast Cancer	Focused on microcalcification
	Kooi T et al (2017)	Breast Cancer	Relatively large scale application
	Geras KJ et al (2017)	Breast Cancer	Multi-view DCNN for high-resolution image

ACKNOWLEDGEMENTS

For technical advice and the drawing in Fig 1., I thank CEO of Lunit Inc., Anthony Seungwook Paek, PhD.

REFERENCES

- Castellino RA. Computer aided detection (CAD): an overview. *Cancer Imaging* 2005;5:17-9.
- Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;175:1828-37.
- He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *CVPR* 2017. arXiv:1512.03385.
- Graves A, Mohamed AR, Hinton G. Speech Recognition with Deep Recurrent Neural Networks. *ICASSP* 2013. arXiv:1303.5778.
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016;316:2402-10.
- Miotto R, Li L, Kidd BA, Dudleya JT. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci Rep* 2016;6:26094.
- Lo SB, Lou S, Lin JS, Freedman MT, Chien MV, Mun SK. Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Transactions on Medical Imaging* 1995;14:711-18.
- Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 1998;86:2278-324.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing System* 2012;1097-105.
- He K, Zhang X, Ren S, Sun S. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition* 2016
- American Cancer Society 2017 Key Statistics of Skin Cancer. Atlanta. <https://www.cancer.org/content/dam/CRC/PDF/Public/8823.00.pdf>
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-18.
- Xu T, Zhang H, Huang X, Zhang S, Metaxas DN. Multimodal Deep Learning for Cervical Dysplasia Diagnosis. *Lecture Notes in Computer Science* 2016;9901:115-23
- Intel & MobileODT Cervical Cancer Screening. <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening>
- American Cancer Society 2017 Key Statistics of Lung Cancer. Atlanta. <https://www.cancer.org/content/dam/CRC/PDF/Public/8703.00.pdf>
- Bar Y, Diamant I, Wolf L, Greenspan H. Deep learning with non-medical training used for chest pathology identification. *Proc. SPIE* 9414, *Medical Imaging 2015: Computer-Aided Diagnosis*, 94140V
- Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, et al. Development of a Digital Image Database for Chest Radiographs

- with and without a Lung Nodule: Receiver Operating Characteristic Analysis of Radiologists' Detection of Pulmonary Nodules. *American Journal of Roentgenology* 2000;174:71-74
18. Bobadilla JCM, Pedrini H. Lung Nodule Classification Based on Deep Convolutional Neural Networks. *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* 2017;117-24
19. Sahiner B, Chan HP, Petrick N, Wei D, Helvie MA, Adler DD, Goodsitt MM. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE Trans Med Imaging*. 1996;15:598-610
20. Jamieson AR, Drukker K, Giger ML. Breast image feature learning with adaptive deconvolutional networks. *Proc. SPIE 8315, Medical Imaging 2012: Computer-Aided Diagnosis*, 831506
21. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesion. *Med Image Anal.* 2017;35:303-12.
22. Hwang S, Kim HE. Self-Transfer Learning for Fully Weakly Supervised Object Localization. arXiv:1602.01625. 23. Wang J, Yang X, Cai H, Tan W, Jin C, Li L. Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. *Sci. Rep.* 2016;6:27327.
24. Geras KJ, Wolfson S, Kim SG, Moy L, Cho K. High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks. arXiv:1703.07047.
25. Samala RK, Chan HP, Hadjiiski L, Helvie MA, Wei J, Cha K. Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography. *Med Phys.* 2016;43:6654
26. Fotin SV, Yin Y, Haldankar H, Hoffmeister JW, Periaswamy S. Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches. *Proc. SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis*, 97850X
27. Zhang Q, Xiao Y, Dai W, Suo J, Wang C, Shi J, Zheng H. Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics* 2016;72:150-57.
28. Dalms MU, Litjens GJ, Holland K, Setio AAA, Mann RM, Karssemeijer N, et al. Using deep learning to segment breast and fibroglandular tissue in MRI volumes. *Medical Physics* 2017;44:533-46.