Effectiveness of the Detection of Pulmonary Emphysema using VGGNet with Low-dose Chest Computed Tomography Images

Doo-Bin Kim[1,2], Young-Joon Park[3], Joo-Wan Hong[4,*]

[1]Korea Medical Institute

[2]Department of Radiology, Eulji University graduate school

[3]Department of Radiologic technology, Cheju halla Univertity

[4]Department of Radiological Science, College of Health Sciences, Eulji University

Received: June 09, 2022. Revised: August 16, 2022. Accepted: August 31, 2022.

ABSTRACT

This study aimed to learn and evaluate the effectiveness of VGGNet in the detection of pulmonary emphysema using low-dose chest computed tomography images. In total, 8000 images with normal findings and 3189 images showing pulmonary emphysema were used. Furthermore, 60%, 24%, and 16% of the normal and emphysema data were randomly assigned to training, validation, and test datasets, respectively, in model learning. VGG16 and VGG19 were used for learning, and the accuracy, loss, confusion matrix, precision, recall, specificity, and F1-score were evaluated. The accuracy and loss for pulmonary emphysema detection of the low-dose chest CT test dataset were 92.35% and 0.21% for VGG16 and 95.88% and 0.09% for VGG19, respectively. The precision, recall, and specificity were 91.60%, 98.36%, and 77.08% for VGG16 and 96.55%, 97.39%, and 92.72% for VGG19, respectively. The F1-scores were 94.86% and 96.97% for VGG16 and VGG19, respectively. Through the above evaluation index, VGG19 is judged to be more useful in detecting pulmonary emphysema. The findings of this study would be useful as basic data for the research on pulmonary emphysema detection models using VGGNet and artificial neural networks.

Keywords: Low-dose chest CT, Emphysema, Convolutional Neural Network, VGGNet

I. INTRODUCTION

Computed tomography(CT) is a basic but important test for disease diagnosis, and the use of low-dose chest CT images is increasing[1]. The rate of early diagnoses of diseases using low-dose chest CT has increased, and various studies have reported its diagnostic usefulness[2,3]. Pulmonary emphysema can be diagnosed using low-dose chest CT. This disease is characterized by inflammation due to the destruction of the alveolar wall in the sub-terminal bronchiolar unit, which subsequently leads to permanently enlarged abnormal alveolar space[4]. Moreover, it is one of the factors causing chronic obstructive pulmonary disease, such as chronic bronchitis, and is a known key risk factor for lung cancer. On CT, pulmonary emphysema presents a unique feature of relatively low attenuation as compared to the normal pulmonary parenchyma[5]. These features are shown in Fig. 1. Recently, the fourth industrial revolution technologies have been used for different purposes and have been proven effective. New technologies such as deep learning have become easily applicable, following the digitalization of medical images such as CT scans[6,7]. A convolutional neural network (CNN), one of the deep learning algorithms, is an artificial

* Corresponding Author: Joo-Wan Hong E-mail: jwhong@eulji.ac.kr Tel: +82-31-740-7380
Address: 553 Sanseong-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do, 13135, Korea

neural network used to analyze images. CNN shows excellent performance in classifying diseases using medical images. Although CheXNet is another model used to detect diseases using radiographic images, it can only detect pneumonia using posteroanterior chest images as the input data[8].

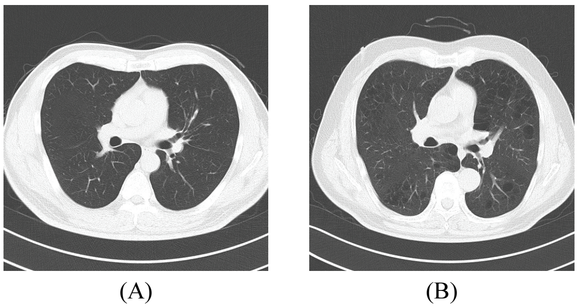


Fig. 1. Low-dose chest CT images, (A) was normal chest CT image and (B) was emphysema chest CT image.

Additionally, various other models using radiographic images are available; however, there is a lack of disease detection models using low-dose chest CT images. Therefore, this study aimed to develop and evaluate a CNN-based VGGNet model through transfer learning for detecting pulmonary emphysema using low-dose chest CT images.

II. MATERIAL AND METHODS

1. Low-dose chest CT imaging data

The data used in this study were collected from patients who underwent low-dose chest CT from January 2020 to January 2021 (IRB approval number: EU22-16). In total, 8000 images with normal findings and 3189 images showing pulmonary emphysema were collected.

2. Data preprocessing

Python (version 3.6) and pycharm integrated development environment (IDE, version 2021.2) were used to develop a pulmonary emphysema detection model. The following libraries were used: tensorflow (version 2.1.0), keras (version 2.3.1), matplotlib (version 3.2.2), and seaborn (0.11.2). Linux, Intel, GeForce, and RAM were used as the learning systems. In total, 11,189 images were used for learning, including 8000 images with normal findings and 3189 images of pulmonary emphysema. For efficient learning, 60%, 24%, and 16% of the normal and emphysema data were randomly assigned for training, validation, and test datasets, respectively. The size of the input data used for learning was converted and fixed to 224 × 224. To increase learning, the original images of the training and validation datasets, excluding the test dataset, were randomly rotated from 0° to 40°, transformed in a counter-clockwise direction within 0.2 radians, moved vertically and horizontally by about 20%, and enlarged and shortened from 80% to 120% in size for image data augmentation.

3. Transfer learning

VGG16 and VGG19 algorithms were used. The general working principle of VGGNet is to apply a 3 × 3 convolutional filter to the convolution layer for extraction of the key features and prevent overfitting through max pooling with 2 × 2 size. Following this, a fully connected layer and flattened layer were used to classify and convert into one-dimensional data. The images were classified based on the features selected through learning[9,10]. Additionally, to improve the learning speed and performance and solve the problems of gradient loss, rectified linear unit(ReLu) activation function was applied to the hidden layer, and a sigmoid activation function was applied for binary classification at the end of the neural network. Adam was used as the optimizer for the optimization of learning[11]. The calculation formula for each function is shown in Eq. (1) ~ (3) below.

$$ReLu = \begin{Bmatrix} x, \; x > 0 \\ 0, \; x \le 0 \end{Bmatrix} = \max(0, x) \tag{1}$$

$$Sigmoid = \frac{1}{1+\exp^{-ax}} = \frac{\exp^{ax}}{1+\exp^{ax}} \tag{2}$$

$$Adam = \theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{\nu}_t} + \epsilon}\widehat{m_t} \tag{3}$$

4. Model evaluation

For the evaluation of the model in which the learning was completed, the accuracy and loss, confusion matrix, precision, recall, specificity, and F1-score for the training and validation data were calculated and evaluated. The calculation formula of each indicator is as shown in (4) ~ (7) below.

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \times 100 \tag{5}$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \tag{6}$$

$$F1\,score = \frac{(2 \times Precision \times Recall)}{(Precision + Recall)} \times 100 \tag{7}$$

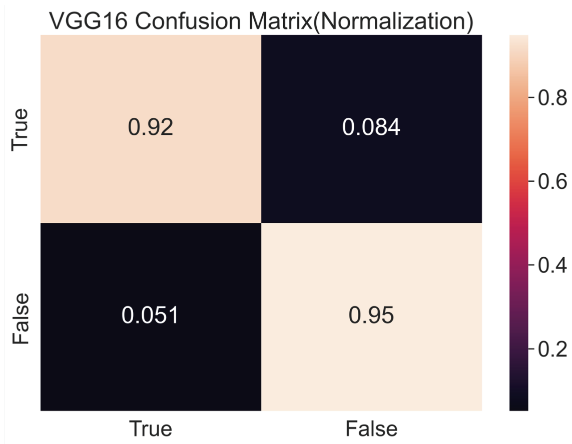
where, TP was true positive and FP was false positive, TN was true negative, FN was false negative.

III. RESULT

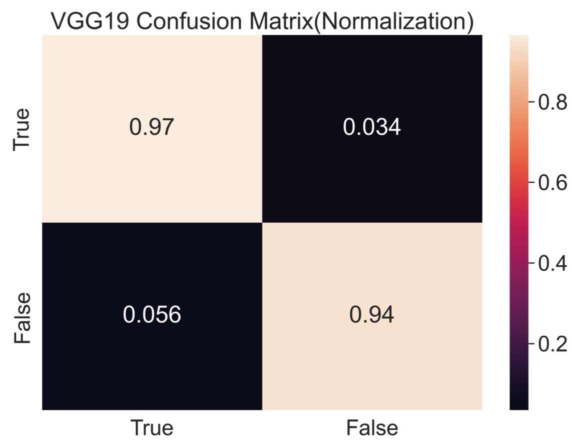
In both learnings using VGG16 and VGG19 models, the training loss change was the smallest at 500 epochs. Thus, subsequent training was stopped. This is represented by Fig. 2. The accuracy and loss for pulmonary emphysema detection of the low-dose chest CT test dataset were 92.35% and 0.21% for VGG16 and 95.88% and 0.09% for VGG19, respectively. The results of the confusion matrix and evaluation indices for the VGG16 and VGG19 models were as follows: precision, 91.60% vs. 96.55%; recall, 98.36% vs. 97.39%; and specificity, 77.08% vs. 92.72%. Additionally, the F1-score was 94.86% and 96.97% for VGG16 and VGG19, respectively. The confusion matrix is shown in the Fig. 3 below, and the results of each evaluation index are shown in the Table 1.

(A) VGG16 (B) VGG19

Fig. 2. Accuracy and loss graphs of VGGNet according to epochs.



(A) VGG16



(B) VGG19

Fig. 3. Confusion matrix of VGGNet.

Table 1. Results of each evaluation index Based on Learning (unit : %)

Evaluation index	VGG16	VGG19
Accuracy	92.35	95.88
Loss	0.21	0.09
Precision	91.60	96.55
Recall	98.36	97.39
Specificity	77.08	92.72
F1-Score	94.86	96.97

IV. DISCUSSION

In this study, VGGNet based on CNN, which has been recently used in medical imaging, underwent learning to detect pulmonary emphysema. The usefulness of the models to diagnose pulmonary emphysema from CT images was assessed. Herein, VGG19 showed superior results than those of VGG16 in all the evaluation indices, except for recall of pulmonary emphysema. This is because VGG19 had a relatively small ratio of False and a high ratio of True when evaluating the model. When comprehensively looking at the evaluation index of this study for the purpose of detecting that the ratio of True should be high, VGG19 with a high ratio of True is considered to be more clinically meaningful. Particularly, the F1-score was greater in VGG19 than in VGG16. However, in this study, pulmonary emphysema was not classified according to the severity. To prevent data bias, various evaluation indicators were used. VGG16 had a relatively low specificity, suggesting that there may have been data bias. The data bias may be solved by increasing the amount of original data. Among many existing image classification models, only VGGNet was used in considering the learning time for each model according to the nature of the processed data. It uses a simple structure and network, which was considered adequate for reducing the physical time[9,12]. Studies using deep learning technology for medical image analysis are being actively conducted. Yao et al. and Wang et al. suggested models for the competent classification of the lesion types[13,14]. However, in this study, significant results were obtained using the existing CNN model rather than other models for competent classification of the lesion types. These results are relevant as they can be used as basic data for studies on CT image analysis using deep learning and for increasing the user accessibility. In another study on the detection of pulmonary emphysema, Braman et al. compared a CNN model using multiple instance learning (MIL) and another model using convolutional-long short term memory (LSTM). The model using LSTM showed superior results when compared to the model using MIL[15]. The LSTM model had 77% recall, 74% specificity, and 75% F1-score, which were relatively lower than those of the model in this study. This difference may be attributed to the level of learning between the models, and it is thought that sufficient learning would not lead to significant differences in using the existing CNN model for the detection of pulmonary emphysema. Bortsova et al. compared measurements of the spatial distribution of emphysema using MIL and the newly proposed learning from label proportions, and Humphries et al. predicted the risk by classifying the severity of pulmonary emphysema[16,17]. These previous findings are relevant for follow-up management of pulmonary emphysema. In this study, the detection rate required for follow-up management was verified. Future studies in combination with the findings of previous studies would be useful for the diagnosis and analysis of pulmonary emphysema using low-dose chest CT images.

V. CONCLUSION

In this study, VGGNet was useful in detecting pulmonary emphysema using low-dose chest CT images. The findings of this study may be used as basic data for studies on customized models for detecting pulmonary emphysema.

Acknowledgement

This work was supported by the Ministry of Educat ion of the Republic of Korea and the National Resear ch Foundation of Korea (NRF-2019R1G1A1010922) a nd this work was received ehical approval from Eulji university IRB approval #EU22-16. This is an IRB-ap proved retrospective study, all patient information was deidentified and patient consent was not required.

Reference

[1] D. J. Brenner, E. J. Hall, "Computed tomography—an increasing source of radiation exposure", New England journal of medicine, Vol. 357, No. 22, pp. 2277-2284, 2007. https://doi.org/10.1056/NEJMra072149

[2] J. C. M. van de Wiel, Y. Wang, D. M. Xu, H. J. van der Zaag-Loonen, E. J. van der Jagt, R. J. van Klaveren, M. Oudkerk, "Neglectable benefit of searching for incidental findings in the Dutch--Belgian lung cancer screening trial (NELSON) using low-dose multidetector CT", European radiology, Vol. 17, No. 6, pp. 1474-1482, 2007. http://dx.doi.org/10.1007/s00330-006-0532-7

[3] P. C. A. Jacobs, W. P. Th. M. Mali, D. E. Grobbee, Y. v. d. Graaf, "Prevalence of incidental findings in computed tomographic screening of the chest: a systematic review", Journal of Computer Assisted Tomography, Vol. 32, No. 2, pp. 214-221, 2008. http://dx.doi.org/10.1097/RCT.0b013e3181585ff2

[4] G. L. Snider, J. Kleinerman, W. M. Thurlbeck, Z. H. Bengali, "The definition of emphysema. Report of a National Heart, Lung, and Blood Institute, Division of Lung Diseases workshop", American Review of Respiratory Disease, Vol. 132, No. 1, pp. 182-185, 1985. https://doi.org/10.1164/arrd.1985.132.1.182

[5] E. J. Stern, M. S. Frank, "CT of the lung in patients with pulmonary emphysema: diagnosis, quantification, and correlation with pathologic and physiologic findings", American Journal of Roentgenology, Vol. 162, No. 4, pp. 791-798, 1994. http://dx.doi.org/10.2214/AJR.162.4.8140992

[6] M. Rastgarpour, J. Shanbehzadeh, "Application of AI Techniques in Medical Image Segmentation and Novel Categorization of Available Methods and Tools", Lecture Notes in Engineering and Computer Science, Vol. 2188, No. 1, pp. 519-523, 2011.

[7] U. Niyaz, A. S. Sambyal, Devanand, "Advances in Deep Learning Techniques for Medical Image Analysis," 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 271-277, 2018. https://doi.org/10.1109/PDGC.2018.8745790

[8] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, A. Y. Ng, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning", arXiv, 2017. https://doi.org/10.48550/arXiv.1711.05225

[9] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image

Recognition", arXiv, 2014. https://doi.org/10.48550/arXiv.1409.1556

[10] X. Glorot, Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks", Journal of Machine Learning Research, Vol. 9, pp. 249-256, 2010.

[11] Y. Bengio, P. Simard, P. Frasconi, "Learning long-term dependencies with gradient descent is difficult", IEEE Transactions on Neural Networks, Vol. 5, No. 2, pp. 157-166, 1994. https://doi.org/10.1109/72.279181

[12] S. H. Kassania, P. H. Kassanib, M. J. Wesolowskic, K. A. Schneidera, R. Detersa, "Automatic Detection of Coronavirus Disease (COVID-19) in X-ray and CT Images: A Machine Learning Based Approach", Biocybernetics and Biomedical Engineering, Vol. 41, No. 3, pp. 867-879, 2021. http://dx.doi.org/10.1016/j.bbe.2021.05.013

[13] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, K. Lyman, "Learning to diagnose from scratch by exploiting dependencies among labels", arXiv, 2017. https://doi.org/10.48550/arXiv.1710.10501

[14] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2097-2106, 2017. https://doi.org/10.1109/CVPR.2017.369

[15] N. Braman, D. Beymer, E. Dehghan, "Disease Detection in Weakly Annotated Volumetric Medical Images using a Convolutional LSTM Network", arXiv, 2018. https://doi.org/10.48550/arXiv.1812.01087

[16] G. Bortsova, F. Dubost, S. Ørting, I. Katramados, L. Hogeweg, L. Thomsen, M. Wille, M. d. Bruijne, "Deep learning from label proportions for emphysema quantification", International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 11071, pp. 768-776, 2018. http://dx.doi.org/10.1007/978-3-030-00934-2_85

[17] S. M. Humphries, A. M. Notary, J. P. Centeno, M. J. Strand, J. D. Crapo, E. K. Silverman, D. A. Lynch, "Deep learning enables automatic classification of emphysema pattern at CT", Radiology, Vol. 294, No. 2, pp. 434-444, 2020. http://dx.doi.org/10.1148/radiol.2019191022

저선량 흉부 CT를 이용한 VGGNet 폐기종 검출 유용성 평가

김두빈[1,2], 박영준[3], 홍주완[4,*]

[1]한국의학연구소
[2]을지대학교 일반대학원 방사선학과
[3]제주한라대학교 보건학부 방사선과
[4]을지대학교 보건과학대학 방사선학과

요 약

본 연구에서는 저선량 흉부 CT 영상을 이용하여 VGGNet을 학습시키고 폐기종 검출 모델을 구현하고 성능을 확인하고자 한다. 연구에 사용된 저선량 흉부 CT 영상은 정상 진단 8000장, 폐기종 진단 3189장이며, 모델 학습을 위해 정상 데이터와 폐기종 데이터를 train, validation, test dataset으로 각각 60%, 24%, 16%로 무작위 추출하여 구분하였다. 학습을 위한 인공신경망은 VGGNet 중 VGG16과 VGG19를 사용하였으며, 학습이 완료된 모델 평가를 위해 정확도, 손실율, 오차 행렬, 정밀도, 재현율, 특이도, F1-score의 평가지표를 사용하였다. 폐기종 검출 정확도와 손실율은 VGG16과 VGG19 각각 92.35%, 95.88%, 0.21%, 0.09%, 정밀도는 91.60%, 96.55%, 재현율은 98.36%, 97.39%, 특이도는 77.08%, 92.72%, F1-score는 94.86%, 96.97%였다. 위의 평가지표를 통해 VGG19 모델의 폐기종 검출 성능이 VGG16 모델에 비해 우수하다고 판단된다. 본 연구를 통해 VGGNet과 인공신경망을 이용한 폐기종 검출 모델 연구에 기초자료로 사용할 수 있을 것으로 사료된다.

중심단어: 저선량 흉부 CT, 폐기종, 합성곱신경망, VGGNet

연구자 정보 이력

	성명	소속	직위
(제1저자)	김두빈	한국의학연구소 을지대학교 일반대학원 방사선학과	보건직 대학원(석사과정)
(공동저자)	박영준	제주한라대학교 보건학부 방사선과	교수
(교신저자)	홍주완	을지대학교 보건과학대학 방사선학과	조교수