

Extraction of Whole-lung Field From Pseudo-chest X-ray Images Using U-net

Shota HOSOGOSHI¹⁾, Kazuaki MATSUO¹⁾, Yoshinori MIYAHARA¹⁾, Yasushi KAJI²⁾

¹⁾ Department of Radiology, Shimane University Hospital, Izumo, Shimane 693-8501, Japan

²⁾ Department of Radiology, Faculty of Medicine, Shimane University, Izumo, Shimane 693-8501, Japan

(Received March 19, 2024; Accepted April 8, 2024; Published online November 22, 2024)

This study aimed to develop a model using U-net to extract the whole lung field from pseudo-chest X-ray images, including areas overlapping with cardiac and diaphragm shadows. Training involved pseudo-X-rays and lung label images from CT scans of 140 cases from the LIDC-IDRI dataset.

The extraction performance of the model was evaluated using the Dice similarity coefficient (DSC). We also examined the correlations among patient size, lung volume, and DSC. As a result, the whole-lung field extraction model developed in this study tended to over-extract intestinal gas in some cases, and the extraction performance varied depending on the patient size. However, the DSC between the whole-lung label image and the output image was >0.9 for all the test data, indicating that the whole-lung field can be extracted from the pseudo chest X-ray image.

Keywords: computed tomography, chest X-ray image, whole-lung field, segmentation, deep learning

Corresponding author: Shota HOSOGOSHI
Department of Radiology, Shimane University Hospital, 89-1
Enya-cho, Izumo, Shimane 693-8501, Japan
Tel: +81-853-20-2435
Fax: +81-853-20-2285
Email: hosogosi@med.shimane-u.ac.jp

INTRODUCTION

Computer-aided diagnosis (CAD) is a software that analyzes and processes medical images to detect abnormalities and performs qualitative diagnosis. CAD in chest X-ray images requires highly accurate lung field extraction because the segmentation of anatomical structures is crucial for detecting lesions in the lung field and measuring the lung area [1]. Several methods have been reported for extracting lung field regions from chest X-ray images using machine learning, including a method based on thresholding using histograms [2], a method that establishes singular points based on anatomical features and uses these singular points [3], and a method for determining the lung field boundaries based on pattern recognition and feature analysis [4].

Recently, deep learning-based methods [5, 6] have been reported for lung field extraction from chest X-ray images, and these methods have achieved higher lung field extraction accuracy than the aforementioned methods. However, as X-ray images depict overlapping structures in the direction of the X-ray projection, it is difficult to extract the whole-lung field from chest X-ray images using conventional lung field extraction methods, including the left lower lobe overlapping the cardiac shadow and bilateral lower lobe lung basement areas overlapping the diaphragm shadow. Extraction of the whole-lung field from chest X-ray images is important because it may contribute in the development of new CAD technology and automatic scan range determination in chest CT examinations; however, only few studies have reported whole-lung field extraction.



This article is licensed under a Creative Commons [Attribution-NonCommercial-NoDerivatives 4.0 International] license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

In this study, we created pseudo-chest X-ray images and whole-lung field label images by processing chest CT images to compose the original dataset. Furthermore, we constructed a whole-lung field extraction model using this dataset and U-Net, an encoder-decoder type fully convolutional network (FCN) [7], as the deep learning model, and verified its extraction performance.

MATERIALS AND METHODS

Collecting image data

The LIDC-IDRI from The Cancer Imaging Archive (TCIA) [8] was used in this study. The LIDC-IDRI database contains data of 1,018 patients who underwent chest CT examinations for lung cancer screening. In this database, the chest CT images of 140 adult patients who met all the following criteria were collected: non-contrast chest CT, slice thickness <3.0 mm, reconstruction kernel as a high-resolution function for the lungs, and no artifacts in the scan area. The collected 140 cases were randomly divided into 100, 20, and 20 groups and used as training, validation, and test data, respectively, to construct the dataset for this study. Training and validation data were used to train the model. During the training process, the training data were used to adjust the model parameters, and the validation data were used to monitor whether the model was overfitting the training data. After training, the performance of the model was evaluated using test data.

Creating pseudo-chest X-ray image

Fig. 1 shows the image processing of a pseudo-chest radiograph. First, the chest CT images (512×512 , 16bit) were loaded into Image J (version 1.48u, National Institutes of Health), adjusted to a window level of -600 and a width of 1600, and then converted to 8bit. Subsequently, coronal images (slice thickness, 3 mm) were created using the re-slice process, and a ray sum image was created by outputting the average pixel values of the coronal images in the anteroposterior direction. Then, to adjust the matrix size of the ray sum image to 512×512 , a margin with a pixel value of 0 was added if the vertical matrix size was <512, and 512 if it was >512. Finally, the images were down sampled to 256×256 pixels. These were defined as pseudo-chest X-ray images in this study. We created pseudo-chest radiographs for all cases collected from the database.

Creating whole-lung field label image

Fig. 2 shows the image processing of a whole-lung field label image. First, the same procedure as for the pseudo-chest X-ray image was used to load the chest CT image into Image J, convert it to 8-bit, and create coronal images. Next, the lung field boundaries of these coronal images were enhanced using a Sobel edge detector. A ray sum image was then created by outputting the average pixel values of the coronal images in the anterior–posterior direction. Binarization was performed for the lung

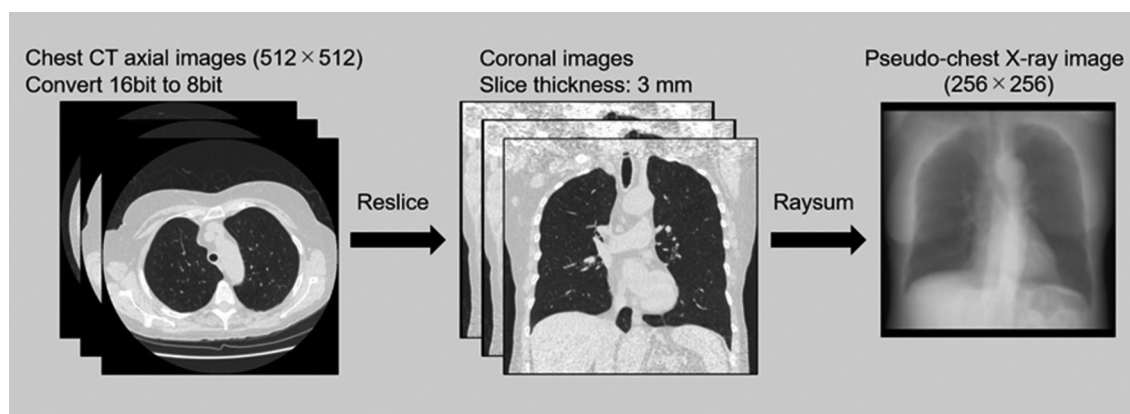


Fig. 1. Creating pseudo-chest X-ray image

The chest CT images were converted to 8 bit. Coronal images were created, and a ray sum image was created by outputting the average pixel values of the coronal images in the anterior-posterior direction. The images were down sampled to 256×256 pixels.

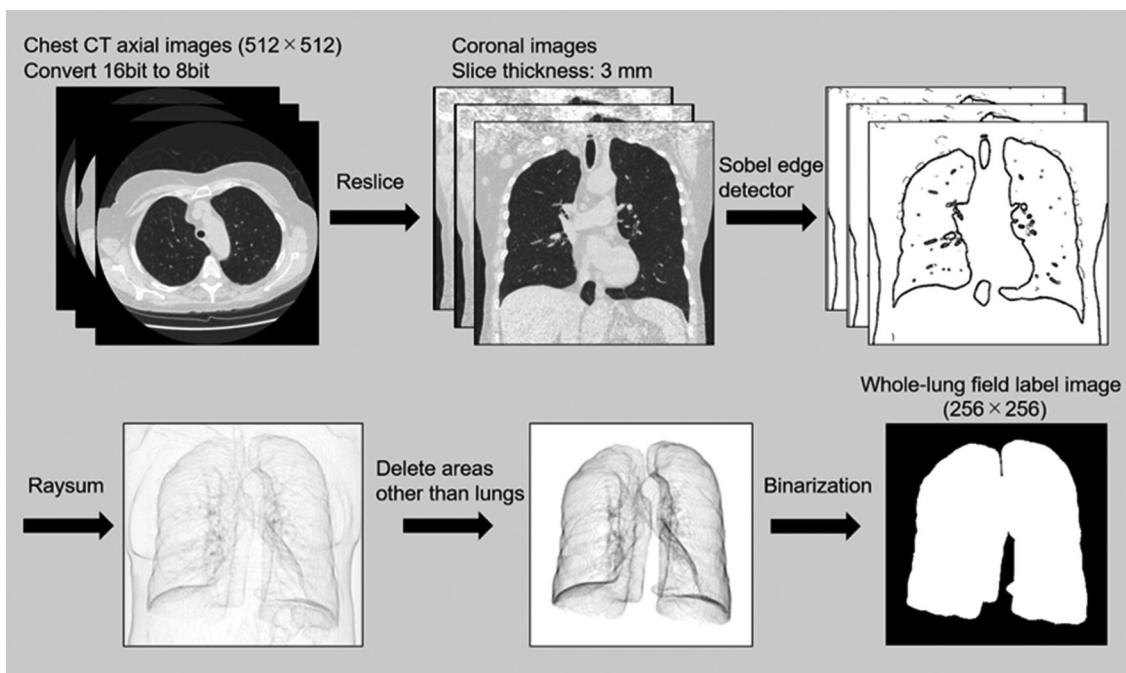


Fig. 2. Creating whole-lung field label image

The chest CT images were converted to 8bit. Coronal images were created, and the lung field boundaries of these coronal images were enhanced using a Sobel edge detector. A ray sum image was created by outputting the average pixel values of the coronal images in the anterior-posterior direction. Binarization was performed for the lung fields and other areas of the images, and areas other than the lungs were deleted. The images were down sampled to 256×256 pixels.

fields and other areas of the image using arithmetic processing, which set the pixel value to 0 for areas that were not lung fields and 255 for lung field areas. Finally, the image size was adjusted using the same procedure as for the pseudo-chest X-ray image. These were defined as whole-lung field-labelled images in this study. We created a whole-lung field label image for all the cases collected from the database.

Extraction of whole-lung fields using U-net

A Neural Network Console (Sony, Tokyo, Japan) was used as the development environment for deep learning models.

The architecture of the U-Net used in this study is shown in Fig. 3. The input image was a pseudo-chest X-ray image and the ground truth was a whole-lung field label image. Training was performed to output the whole lung field, including the left lower lobe overlapping the cardiac shadow of the pseudo-chest X-ray image and the bilateral lower lobe lung basement areas overlapping the diaphragmatic shadow. Binarization of the lung and non-lung fields was performed by setting the threshold of the

pixel value of the lung fields in the output image to 0.5. Binary cross-entropy was used as the loss function during training and the model with the smallest loss was adopted as the extraction model. The learning parameters were set as follows: number of epochs, 50; batch size, 5; Adam optimization function; and learning rate, 0.001. The holdout method was used to evaluate the extraction model.

Evaluation method

In this study, the Dice similarity coefficient (DSC) was used to evaluate the similarity of the lung fields between the whole-lung field label image and the output image from the extraction model in 20 test data cases.

$$\text{Dice similarity coefficient}_{(X,Y)} = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (1),$$

where X is the number of lung field pixels in the whole-lung field label image, and Y is the number of lung field pixels in the output image. Furthermore, we focused on the correlation among patient size, lung volume (LV), and DSC. As an index of patient size, effective diameter (ED) [9] was calculated from the anterior-posterior and lateral lengths

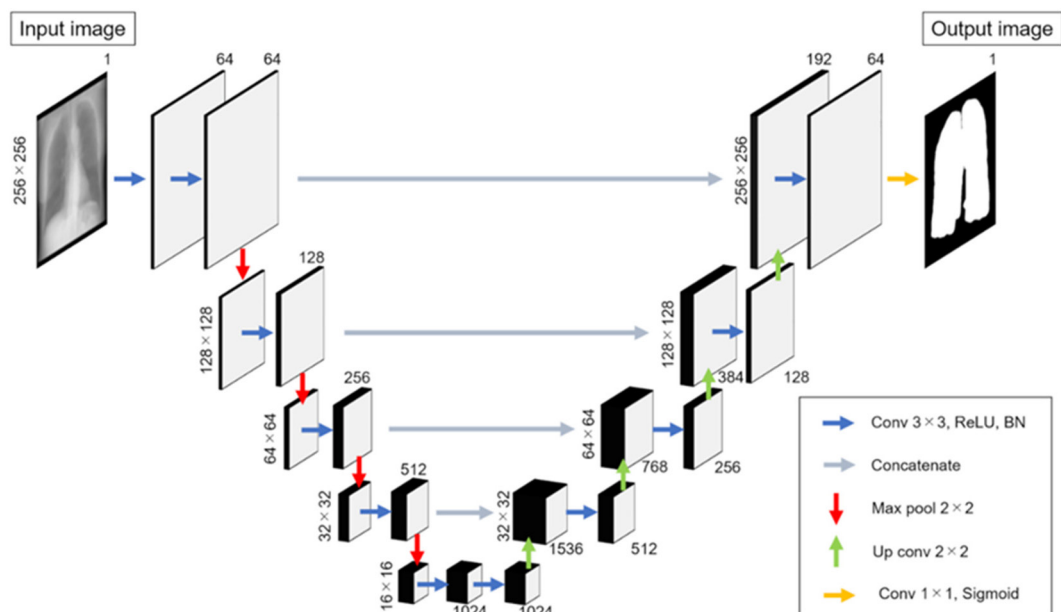


Fig. 3. Architecture of U-net in this study

The encoder part of U-Net consists of five stages, and at each stage, convolution, batch normalization, and max pooling are repeated to the bottom of the U-Net. The number of convolution filters was increased to 64, 128, 256, 512, and 1,024. In the convolution layer, a filter with a kernel size of 3×3 was convolved with stride 1, and a rectified linear unit was used as the activation function. At the bottom of the U-Net, convolution and batch normalization were performed twice and output to the decoder section. In the decoder section, the output from the encoder section of the same resolution was concatenated, and convolution, batch normalization, and up conversion were repeated. In the last convolution layer, a filter with a kernel size of 1×1 was convolved with stride 1, and a sigmoid function was used to output a 256×256 -pixel image.

measured from the axial images in the lung base-ment area of each test data.

$$ED = \sqrt{AP \times LAT} \quad (2),$$

where AP is the anteroposterior length and LAT is the lateral length. The LV of the test data was measured using a 3D workstation (SYNAPSE VIN-CENT; FUJIFILM, Tokyo, Japan). The correlation among ED, LV, and DSC was examined using the single regression analysis.

Statistical analysis

Numerical data were expressed as the mean \pm stan-dard deviation. The Spearman's rank correlation coefficient was used to test for correlations. The sta-tistical significance level was set at 5%. R (version 4.0.4, The R Foundation for Statistical Computing) [10] was used for statistical analysis.

RESULTS

Evaluation of similarity

In the 20 test data cases, ED was 293.47 ± 30.21

mm, LV was 4579.92 ± 970.55 ml, and DSC be-tween whole-lung field label image and output im-age was 0.966 ± 0.015 (Table 1). The DSC was >0.9 for all the test data; therefore, the whole-lung field label image and the output image were in good agreement.

Fig. 4 shows five examples of the lung field ex-traction results. In this figure, from left to right are the pseudo-chest X-ray, whole-lung label, and output images. In all the test data, (a) is the case with the smallest ED, (b) is the case with the largest ED, (c) is the case with the smallest LV, and (d) is the case with the largest LV. The DSCs were 0.974, 0.960, 0.946, and 0.984 for (a), (b), (c), and (d), respectively, (e) shows a case of the overextraction of intestinal gas. The DSC in this case was the lowest for all the test data, with a value of 0.928.

Correlation among patient size, LV, and DSC

Fig. 5 shows the correlation between ED and DSC, and between LV and DSC. There was a significant negative correlation between ED and DSC ($r =$

Table 1. Effective diameter, Lung volume and Dice similarity coefficient in the 20 test data cases

Effective diameter (mm)	Lung volume (ml)	Dice similarity coefficient
293.47 ± 30.21	4579.92 ± 970.55	0.966 ± 0.015

Data are the mean \pm standard deviation

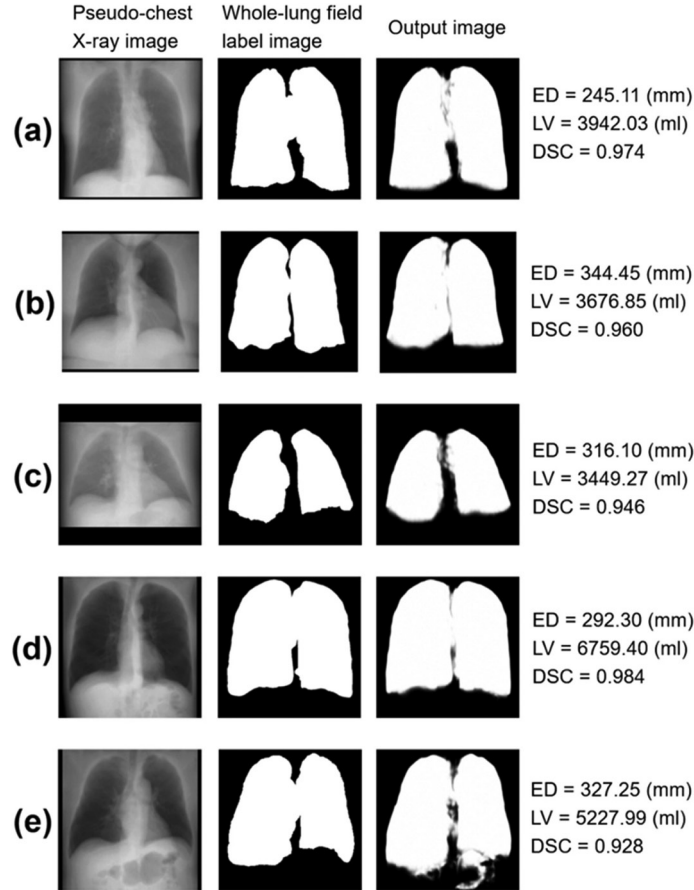


Fig. 4. Examples of lung field extraction results

(a) case with the smallest ED, (b) case with the largest ED, (c) case with the smallest LV, (d) case with the largest LV, (e) case of over-extraction of intestinal gas.
 ED: effective diameter; LV: lung volume; DSC: Dice similarity coefficient

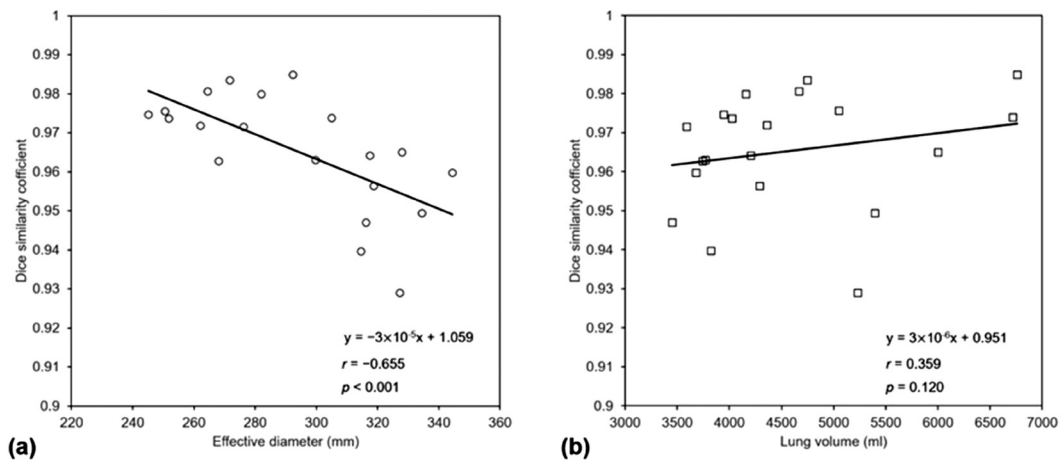


Fig. 5. Correlation among patient size, lung volume, and DSC

(a) correlation between ED and DSC, (b) correlation between LV and DSC.
 ED: effective diameter; LV: lung volume; DSC: Dice similarity coefficient

-0.655, $p < 0.001$). There was a weak positive but insignificant correlation between LV and DSC ($r = 0.359$, $p = 0.120$).

DISCUSSION

The whole-lung field extraction model in this study was able to extract the left lower lobe overlapping the cardiac shadow and the bilateral lower lobe lung basement areas overlapping the diaphragmatic shadow in most test data. Furthermore, the DSC between the whole-lung label image and the output image was >0.9 for all the test data. Gozes *et al.* reported a deep-learning-based image processing technique for enhancing the contrast of soft lung structures in chest X-ray images using a FCN [11]. In the process of accomplishing this task, they created pseudo-chest X-ray images, labeled images of the whole-lung field from chest CT images, and used these images to perform deep learning to extract the whole-lung field. Comparing their whole-lung field extraction results with ours, their model had a DSC of 0.953, while our model had a DSC of 0.966 ± 0.015 , indicating comparable extraction performance. U-net is an encoder–decoder-type FCN that does not have a full concatenation layer; however, it consists of a convolutional layer [7]. The feature map in the convolution contains simple and concrete features of the image as the layers of the network become shallower, and complex and abstract features of the image as the layers become deeper. The Fourier transform of the feature map in the convolution is a high-pass filter that amplifies high-frequency components [12]. Therefore, the high extraction performance of the whole-lung field in U-net and FCN can be attributed to the fact that the left lower lobe overlaps the cardiac shadow, and the bilateral lower lobe overlaps the diaphragm shadow, reducing the ambiguity at the boundary in the lung basement area by adding global and local feature maps of the lungs through skip connections.

In contrast, case (e) in Fig. 4 shows an overextraction of intestinal gas, with a DSC of 0.928, the lowest value in the test case. Factors contributing to the overextraction of intestinal gas may be the small amount of training data where intestinal gas was present and the structure of the extraction

model. Collecting and learning from many cases in which intestinal gas is present may reduce the overextraction of intestinal gas. Furthermore, changing the model from U-net to Mask R-CNN also reduced the overextraction of intestinal gas. Mask R-CNN obtains the feature map of the image using a CNN, inputs the feature map to the regional proposal network, and detects the location and content in the image [13]. Then, even if they are in the same class and are different individuals, instance segmentation recognizes them as different individuals and performs segmentation. Therefore, the Mask R-CNN can learn to recognize the lung field and intestinal gas as separate individuals, which is expected to improve the performance of the whole-lung field extraction models.

As for the correlation among patient size, LV, and DSC, there was a significant negative correlation between ED and DSC ($r = -0.655$, $p < 0.001$), as shown in Fig. 5. People with high visceral fat have diaphragm elevation because caudal diaphragm movement is impeded compared with those who do not have high visceral fat. This may have caused an increase in the overlap between the bilateral lower lobe lung basement areas and the diaphragm on the chest radiograph, resulting in a tendency for the DSC to be lower in cases with larger EDs because of the lower extraction performance of the bilateral lower lobe lung basement areas. In contrast, there was a weak positive but insignificant correlation between LV and DSC ($r = 0.359$, $p = 0.120$). The results showed that LV did not affect the extraction performance of the whole-lung field.

Most previous CAD systems for lung disease detection in chest X-ray images performed lung field extraction before detecting the disease, and then detected the disease based on the extracted regions [14]. These systems are unable to extract the left lower lobe, which overlaps with the cardiac shadow, and the bilateral lower lobe lung basement areas, which overlap with the diaphragmatic shadow, thus detecting the disease in a limited lung area. However, our proposed method enables the detection of lung diseases, including areas that could not be extracted previously, and will contribute to the development of novel CAD technologies. In addition, Demircioğlu *et al.* used conditional generative ad-

versarial networks to achieve automatic scan range setting for chest CT [15]. The automatic scan range setting reduces a patient's radiation dose by suppressing overscanning and reduces the variation in range setting among radiological technologists [16]. In the study by Demircioğlu *et al.* the radiologist wrote annotations for the start and end positions of the scan in the scout view, and the neural network learned the positions of the annotations. Subsequently, an image is generated with annotations of the start and end positions of the scan for an unknown scout view. However, this approach may not consider anatomical structures during the learning process. Additionally, the usefulness of the model should be determined under the supervision of a radiologist. In contrast, our proposed method can set the scan range based on the anatomical structure by extracting the whole-lung field with high accuracy, which contributes to improving the reliability of automatic scan range setting.

One limitation of this study was that it was performed on cases with clear lung field boundaries and did not include cases in which the disease obscured the boundaries of the lung fields. In patients with disease, lung areas are reported to be difficult to segment due to stiffening, cloudiness, cavities, and masses in the lungs [17]. Therefore, future studies should include additional cases with unclear lung field boundaries and examine the robustness of the model in cases involving various diseases. In addition, the hyperparameters of the whole-lung field extraction model were not examined in this study. For example, batch normalization was used for the model in this study; however, the batch size was set to five owing to GPU memory capacity limitations. In deep learning models that use batch normalization, if the batch size is too small, the estimation of the mean and variance during the normalization of the neural network activity values may become unstable, and learning may not proceed well [18]. As the model in this study obtained high extraction performance even with a batch size of five, we did not verify the model by varying the batch size. However, it is necessary to consider hyperparameters, such as batch size, to further improve the performance of this model.

Since the whole-lung field extraction model in

this study uses pseudo-chest X-ray images as training data, the model may need to be adjusted for real-world patient data. Therefore, our future plan is to validate the performance of the whole-lung field extraction on real-world patient data and to adjust improve the performance of the model.

CONCLUSIONS

The whole-lung field extraction model developed in this study tended to over-extract intestinal gas in some cases, and the extraction performance varied depending on patient size. However, the DSC between the whole lung label image and the output image was >0.9 for all the test data, indicating that the whole-lung field can be extracted from the pseudo-chest X-ray image.

Ethical approval

All procedures performed in the current study involving human participants were in accordance with the ethical standards of the Institutional Review Board (IRB) and the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Our study used anonymized CT scans from the LIDC-IDRI database. This database is based on the clinical data collected from medical institutions and is publicly available for research purposes. Formal informed consent was not required for this type of study at our institution.

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Shota Hosogoshi and Kazuaki Matsuo. The first draft of the manuscript was written by Shota Hosogoshi and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

None.

Funding

None.

Conflicts of Interests

All authors have no conflicts of interests to declare.

REFERENCES

- 1) Dai W, Doyle J, Liang X, *et al.* Scan: Structure Correcting Adversarial Network for Organ Segmentation in Chest X-Rays. In: Stoyanov D, Taylor Z, Carneiro G, *et al.* *Deep learning in medical image analysis and Multimodal learning for clinical decision support, DLMLA 2018*. Berlin: Springer; 2018:263-273. doi: 10.1007/978-3-030-00889-5_30.
- 2) Armato SG 3rd, Giger ML, MacMahon H. Automated lung segmentation in digitized posteroanterior chest radiographs. *Acad Radiol.* 1998;5:245-255. doi: 10.1016/s1076-6332(98)80223-7.
- 3) Brown MS, Wilson LS, Doust BD, Gill RW, Sun C. Knowledge-based method for segmentation and analysis of lung boundaries in chest X-ray images. *Comput Med Imaging Graph.* 1998;22:463-477. doi: 10.1016/S0895-6111(98)00051-2.
- 4) Li L, Zheng Y, Kallergi M, Clark RA. Improved method for automatic identification of lung regions on chest radiographs. *Acad Radiol.* 2001;8:629-638. doi: 10.1016/S1076-6332(03)80688-8.
- 5) Rashid R, Akram MU, Hassan T. Fully Convolutional Neural Network for Lungs Segmentation from Chest X-Rays. In: Campilho A, Karray F, Romeny BT. *Image Analysis and Recognition (ICIAR 2018)*. Vol. 10882. Berlin: Springer; 2018:71-80. doi: 10.1007/978-3-319-93000-8_9.
- 6) Islam J, Zhang Y. Towards robust lung segmentation in chest radiographs with deep learning. arXiv:1811.12638 [Preprint]. 2018 [cited 2018 Nov 30]: [5 p]. Available from: <https://arxiv.org/abs/1811.12638>. doi: 10.48550/arXiv.1811.12638
- 7) Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF. *MICCAI, PT III*. vol. 9351. Berlin: Springer; 2015:234-241. doi: 10.1007/978-3-319-24574-4_28.
- 8) Clark K, Vendt B, Smith K, *et al.* The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging.* 2013;26(6):1045-57. doi: 10.1007/s10278-013-9622-7.
- 9) American Association of Physicists in Medicine. Size-specific dose estimates (SSDE) in pediatric and adult body CT examinations (task Group 204). American Association of Physicists in Medicine. https://www.aapm.org/pubs/reports/rpt_204.pdf. (accessed 28 Nov 2022).
- 10) R Core Team. R: A Language and Environment for Statistical Computing; 2021. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>. (accessed 28 Nov 2022).
- 11) Gozes O, Greenspan H. Lung structures enhancement in chest radiographs via CT based FCNN training. In: Stoyanov D, Taylor Z, Kainz B, Maicas G, Beichel RR. *Image analysis for moving organ, breast, and thoracic images*. Berlin: Springer; 2018:147-158. doi: 10.1007/978-3-030-00946-5_16.
- 12) Park N, Kim S. How do vision transformers work?. arXiv:2202.06709 [Preprint]. 2022 [cited 2022 Jun 8]: [26 p]. Available from: <https://arxiv.org/abs/2202.06709>. doi: 10.48550/arXiv.2202.06709.
- 13) He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. arXiv:1703.06870 [Preprint]. 2017 [cited 2018 Jan 24]: [12 p]. Available from: <https://arxiv.org/abs/1703.06870>. doi: 10.48550/arXiv.1703.06870.
- 14) Zaidi SZY, Akram MU, Jameel A, Alghamdi NS. Lung segmentation-based pulmonary disease classification using deep neural networks. In: *IEEE Access*. Vol. 9. Piscataway: IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC; 2021:125202-125214. doi: 10.1109/ACCESS.2021.3110904.
- 15) Demircioğlu A, Kim MS, Stein MC, *et al.* Automatic scan range delimitation in chest CT using deep learning. *Radiol Artif Intell.* 2021;3(3):e200211. doi: 10.1148/ryai.2021200211.
- 16) Colevray M, Tatard-Leitman VM, Gouttard S, Douek P, Bousset L. Convolutional neural network evaluation of over-scanning in lung computed tomography. *Diagn Interv Imaging.* 2019;100:177-183. doi: 10.1016/j.diii.2018.11.001.
- 17) Dallal AH, Agarwal C, Arbabshirani MR, Patel A, Moore G. Automatic estimation of

- heart boundaries and cardiothoracic ratio from chest X-ray images. *Proceedings of SPIE* 2017;10134:134-143. doi: 10.1117/12.2254136
- 18) Lian X, Liu J. Revisit batch normalization: New understanding and refinement via composition optimization. In: Chaudhuri K, Sugiyama M. *The 22nd International Conference on Artificial Intelligence and Statistics*. Brookline, MA: Microtome Publishing; 2019:3254-3263.

