Original Article

Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement without self-annotation

Kotaro Yoshida[1]^, Atsushi Takamatsu[1], Takashi Matsubara[1], Taichi Kitagawa[1], Fomihito Toshima[1], Rie Tanaka[2], Toshifumi Gabata[1]

[1]Department of Radiology, Kanazawa University Graduate School of Medical Sciences, Kanazawa, Japan; [2]College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, Kanazawa, Japan

Contributions: (I) Conception and design: K Yoshida, A Takamatsu; (II) Administrative support: R Tanaka, T Gabata; (III) Provision of study materials or patients: K Yoshida, A Takamatsu; (IV) Collection and assembly of data: K Yoshida, A Takamatsu; (V) Data analysis and interpretation: K Yoshida, A Takamatsu, T Matsubara, T Kitagawa, R Tanaka; (VI) Manuscript writing: All authors; (VII) Final approval of the manuscript: All authors.

Correspondence to: Kotaro Yoshida, MD, PhD. Department of Radiology, Kanazawa University Graduate School of Medical Sciences, 13 1 Takaramachi, Kanazawa, Ishikawa 920 8641, Japan. Email: kotaro_yoshida@staff.kanazawa-u.ac.jp.

Background: A reproducible and accurate automated approach to measuring cardiothoracic ratio on chest radiographs is warranted. This study aimed to develop a deep learning-based model for estimating the cardiothoracic ratio on chest radiographs without requiring self-annotation and to compare its results with those of manual measurements.

Methods: The U-net architecture was designed to segment the right and left lungs and the cardiac shadow, from chest radiographs. The cardiothoracic ratio was then calculated using these labels by a mathematical algorithm. The initial model of deep learning-based cardiothoracic ratio measurement was developed using open-source 247 chest radiographs that had already been annotated. The advanced model was developed using a training dataset of 729 original chest radiographs, the labels of which were generated by the initial model and then screened. The cardiothoracic ratio of the two models was estimated in an independent test set of 120 original cases, and the results were compared to those obtained through manual measurement by four radiologists and the image-reading reports.

Results: The means and standard deviations of the cardiothoracic ratio were 52.4% and 9.8% for the initial model, 51.0% and 9.3% for the advanced model, and 49.8% and 9.4% for the total of four manual measurements, respectively. The intraclass correlation coefficients (ICCs) of the cardiothoracic ratio ranged from 0.91 to 0.93 between the advanced model and the manual measurements, whereas those for the initial model and the manual measurements ranged from 0.77 to 0.82.

Conclusions: Deep learning-based cardiothoracic ratio estimation on chest radiographs correlated favorably with the results obtained through manual measurements by radiologists. When the model was trained on additional local images generated by the initial model, the correlation with manual measurement improved even more than the initial model alone.

Keywords: Deep learning; cardiothoracic ratio measurement; chest radiograph

Submitted Feb 15, 2023. Accepted for publication Jul 21, 2023. Published online Sep 13, 2023.
doi: 10.21037/qims-23-187
View this article at: https://dx.doi.org/10.21037/qims-23-187

^ ORCID: 0000-0002-8398-0969.

Introduction

The cardiothoracic ratio is a value calculated from a chest radiograph and is used to assess the size of the heart relative to the chest cavity (1). This ratio can vary based on factors such as the individual's body size, imaging conditions, and degree of inspiration during the radiograph. The presence of cardiomegaly, or an enlarged heart, on a chest radiograph can be used to screen for various diseases such as cardiac conditions and mediastinal lesions (2,3). In countries where hemodialysis is readily available, the cardiothoracic ratio is also used as an indicator to maintain the appropriate dry weight in dialysis patients (4).

In current clinical practice, the cardiothoracic ratio is manually measured and documented by the radiologist or attending physician. However, this approach often results in inconsistencies in inter-reader and intra-reader agreement (5). To address this problem, various automatic measurement methods for the cardiothoracic ratio have been developed. However, most of those methods have not been widely adopted, and there are still several issues that need to be addressed to make these methods suitable for use in daily practice, such as accuracy and versatility issues (6). Recently, there has been increased interest in developing an automatic measurement model using deep learning (5). While it is anticipated that deep learning-based models will be highly accurate, such models require a large number of training datasets with annotation or labeling to increase their accuracy (7). While there has been a demand for the development of models that minimize the effort required to create training datasets, there have been few reports on this topic to date (8). In this study, we developed a deep learning-based model using open-source training data, and then expanded the training data with local data to create a more advanced model. The purpose of this study was to develop a deep learning-based model for estimating the cardiothoracic ratio using U-net, without the need for additional manual annotation. We also compared the results of these models with those of manual measurements.

Methods

This retrospective study was conducted at Kanazawa University Hospital and was approved by the institutional review board of this hospital. The requirement for informed consent was waived due to the retrospective nature of the study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Chest radiographs of 247 cases from open-source datasets and 1,495 cases from Kanazawa University Hospital were used to create a training dataset for the development of a deep learning-based model. The 247 open-source cases were obtained from the miniJSRT_database provided by the Japanese Society of Radiological Technology (9). The images consist of the original image in the BMP file format with a matrix size of 256×256, and a label image in PNG file format with pixel values 255 for lung, 85 for mediastinum, 170 for non-lung area, and 0 for the area outside the body. The images of 1,495 chest radiographs collected from our institutional image database were acquired between January 2020 and December 2020 with the cardiothoracic ratio noted in the imaging report. The 120 images in the test set included 30 randomly selected patients in each of four groups with cardiothoracic ratios of <40%, 40–49%, 50–59%, and ≥60% based on the cardiothoracic ratio in imaging reports from January to June 2021.

Two types of deep learning-based models were developed based on the U-net architecture. U-net is a popular architecture for image segmentation that consists of U-shaped encoder-decoder network (10). Our architecture of U-net is shown in *Figure 1*. The detailed parameters for our U-net model are shown in Table S1. The same U-net architecture was used for both the initial and advanced models. All U-net processing was performed on a Windows 10 (×64) workstation with two NVIDIA GeForce 2080Ti GPUs with 24 GB memory and were implemented with Python (version 3.8.8) and PyTorch (version 1.1.0). We trained the models using the Adam optimizer with a minibatch size of 64 for 1,000 epochs and a learning rate of 0.0001. The image data were converted into a BMP file with a matrix size of 256×256 for the input image, and the output image was a PNG file with a matrix size of 256×256 with pixel values 255 for left lung, 170 for right lung, 85 for mediastinum, and 0 for the area outside the body. The initial model was created using the 247 open-source images with segmentation information of the right lung, left lung, and heart. To develop the advanced model, additional training data were prepared by segmentation labels on new images and were screened. In total, 1,495 original images were segmented using the initial model. The images with contour information of segmentation for both lungs and the heart were visually evaluated by two certified radiologists (with 17- and 14-year experience) to select an appropriate training dataset. The images were classified into three groups: excellent, good, and fair. Excellent was defined

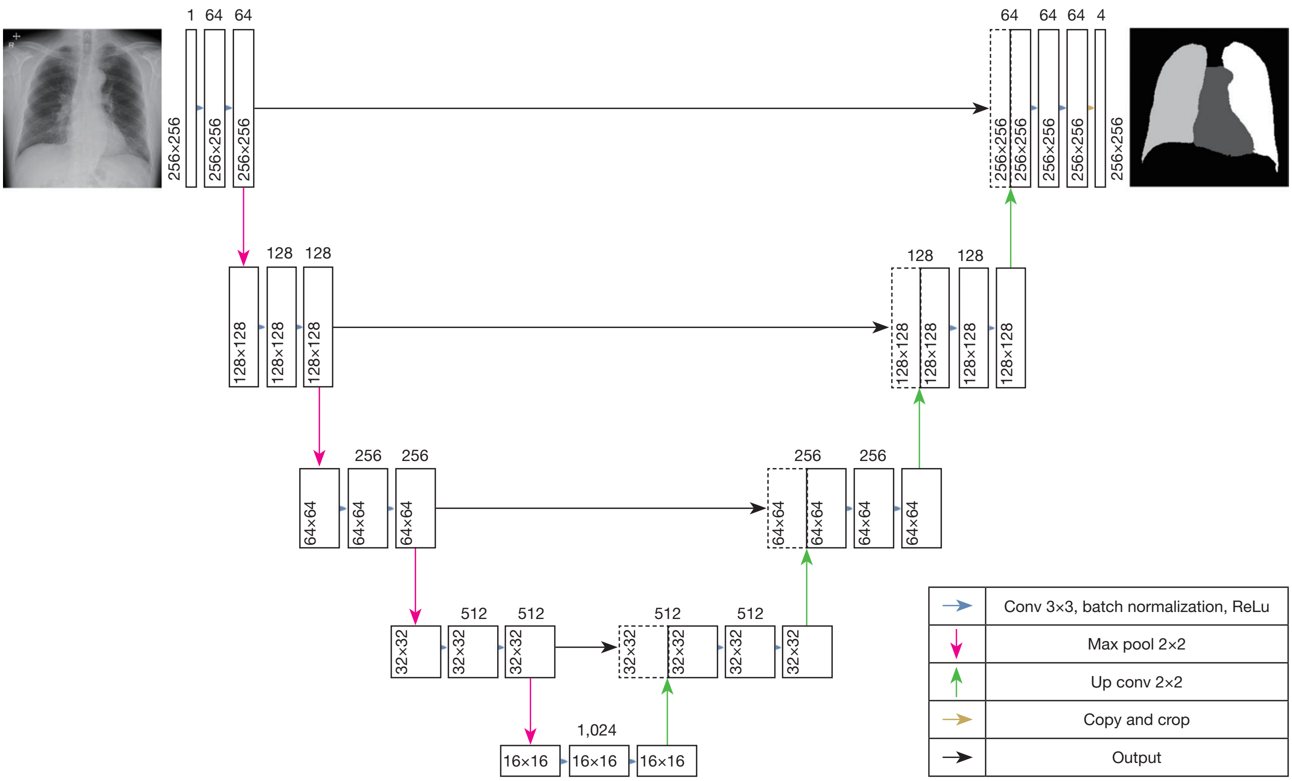


Figure 1 A U-net architecture for segmentation of lung and heart from a chest radiograph.

as good segmentation of the heart and lungs, suitable for measuring the cardiothoracic ratio; good was defined as some defective segmentation of the heart and lungs, which may have a minor effect on measuring the cardiothoracic ratio; and fair was defined as poor segmentation of the heart and lungs, which was not suitable for measuring the cardiothoracic ratio. A flowchart of model development, cardiothoracic ratio measurement output, and an example of segmentation is shown in Figures S1-S3. The cases that were deemed excellent by both two readers were used as a new training dataset for the advanced model. Ultimately, 729 chest radiographs were included in the training dataset for the advanced model. With this training dataset, we developed an advanced deep learning-based model with semantic segmentation using U-net, in the same manner as the initial model.

The cardiothoracic ratio was calculated by dividing the transverse axis length of the right margin of the segmented right lung and the left margin of the left lung (representing the thorax) by the transverse axis length of the right and left margins of the heart (representing the heart). Image segmentation and subsequent sequential cardiothoracic ratio measurements were performed on the independent test dataset using the two models, and cardiothoracic ratios were output for each of the 120 test cases.

The reference for the cardiothoracic ratio was measured by four radiologists (with 18, 15, 5 and 3 years of experience), and the description of cardiothoracic ratio in the radiology report was also used as a reference. The measurement of the cardiothoracic ratio was performed using a medical image viewer (EV Insite, PSP Co., Tokyo, Japan). The results of automatic measurement of the cardiothoracic ratio on the test dataset by the initial and advanced models were compared with the results of manual measurement.

The cardiothoracic ratios estimated by two deep learning-based models were compared by the intraclass correlation coefficient (ICC), and the errors between the two models and the radiologist's manual measurements were examined by a Bland-Altman plot. R version 4.0.3 was used for the graphing and statistical processing. To evaluate the segmentation performance of the two models, we calculated the Dice similarity coefficient (DSC) between each model and the manual segmented data by a radiologist.

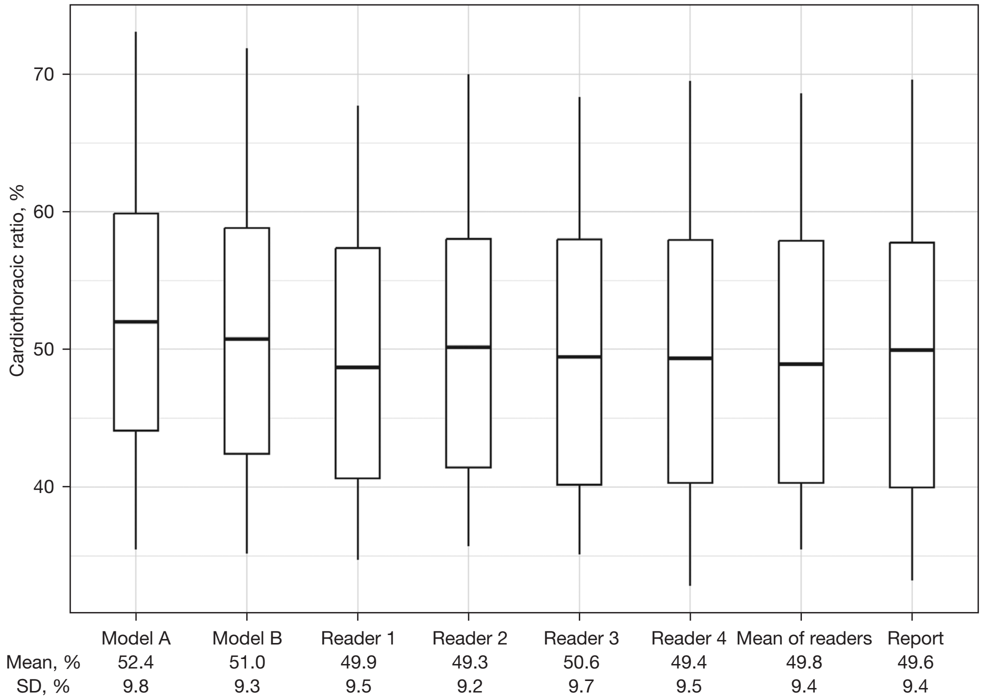


Figure 2 The cardiothoracic ratio in the deep learning-based models and the references in the test set. The box plot shows the medians and first and third quartiles in boxes, as well as the minimum and maximum values of the cardiothoracic ratio for the deep-learning models and references. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists; Report, cardiothoracic ratio measured by radiology reports; SD, standard deviation.

We randomly selected 40 patients from the test set for this evaluation.

Results

The cardiothoracic ratio in the artificial intelligence (AI) model, manual readings, and reading reports in the test data are shown in *Figures 2,3*. The mean cardiothoracic ratio was slightly larger in the AI model than in the manual readings and reading reports (52.4% and 9.8% for the initial model, 51.0% and 9.3% for the advanced model, and 49.8% and of 9.4% for the total of four manual measurements). Correlation analysis of the cardiothoracic ratio in each case showed a higher correlation with manual measurements in the advanced model than in the initial model (ICC: 0.91 to 0.93 between the advanced model and the manual measurements, ICC: 0.77 to 0.82 between the initial model and the manual measurements).

The relationship between each AI model and manual readings/reading reports was examined using Bland-Altman plots (*Figure 4*). The advanced model showed a decrease in error compared to the initial model. No systematic errors were found for either model. Examples of segmentation by the initial and advanced models are shown in *Figures 5,6*. The DSCs between each model and the manual segmentation were follows: for the initial model, heart: mean 0.96, standard deviation (SD) 0.02, lung: mean 0.98, SD 0.01; for the advanced model, heart: mean 0.97, SD 0.02, lung: mean 0.99, SD 0.01.

Discussion

In this study, we attempted to develop deep learning-based models to estimate the cardiothoracic ratio in chest radiographs. The advanced model, which was trained using data acquired without additional segmentation, showed strong correlation with the reference values of cardiothoracic ratio obtained through manual measurement. The advanced model had an ICC of approximately 0.93 with manual measurements, indicating that it could

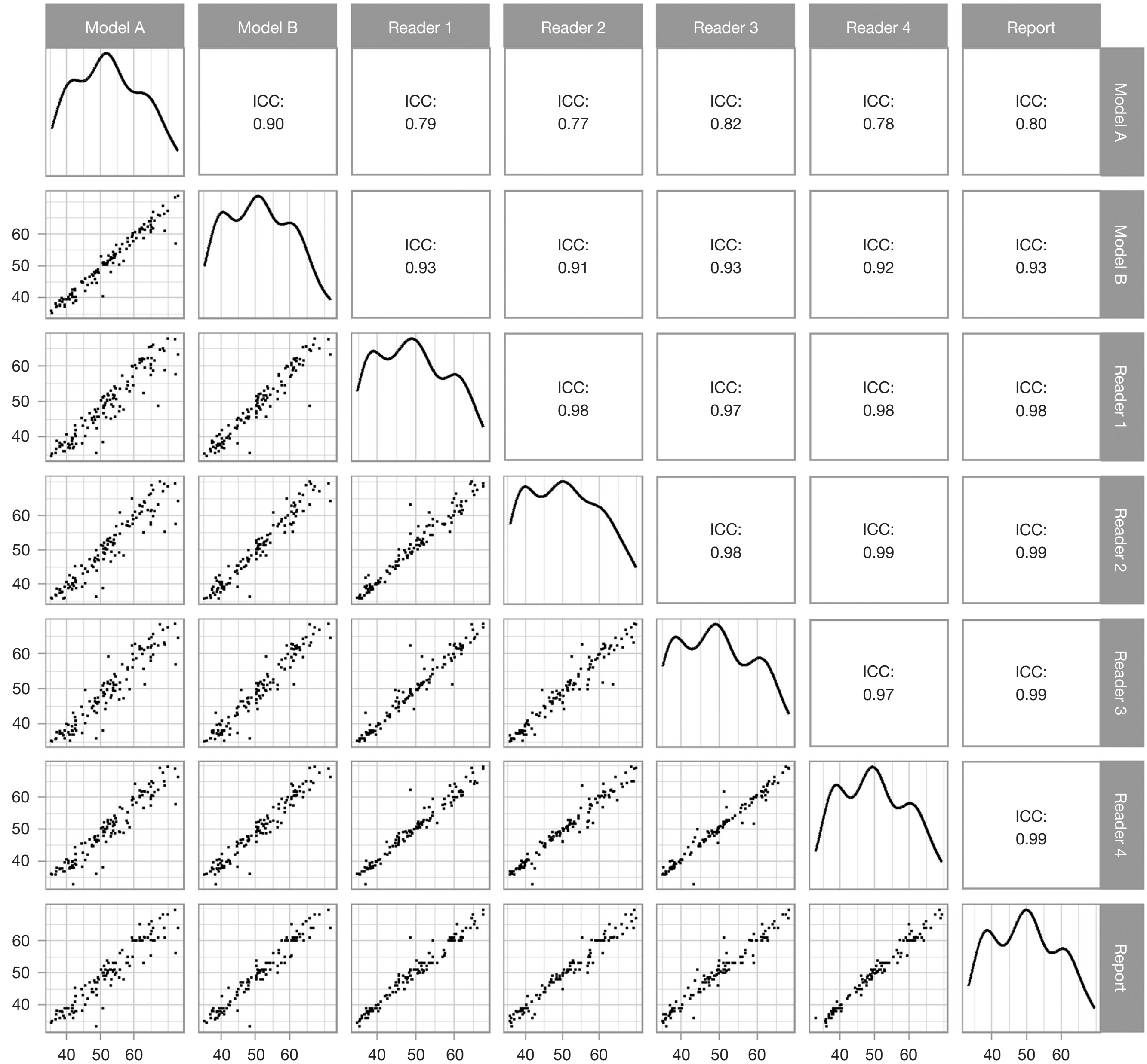


Figure 3 Correlation of the cardiothoracic ratio between the two deep learning models, manual reading by four radiologists, and radiology report in the test dataset. The correlation matrix shows the relationship between the two deep-learning models, radiology report, and manual measurements by four radiologists. The X-axis represents the cardiothoracic ratio (%) calculated by the upper item, and the Y-axis represents the cardiothoracic ratio (%) calculated by the right-hand item. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists; Report, cardiothoracic ratio measured by radiology reports; ICC, intraclass correlation coefficient.

potentially be used in clinical practice.

Measurements of the target lesion in diagnostic imaging including abnormal lesions and normal anatomical structure are used in a variety of situations. The measurement of abnormal lesions is important in diagnostic imaging because it is related to disease progression and severity. However, until recently, the measurement of normal anatomical structures was not considered significant. Longitudinal

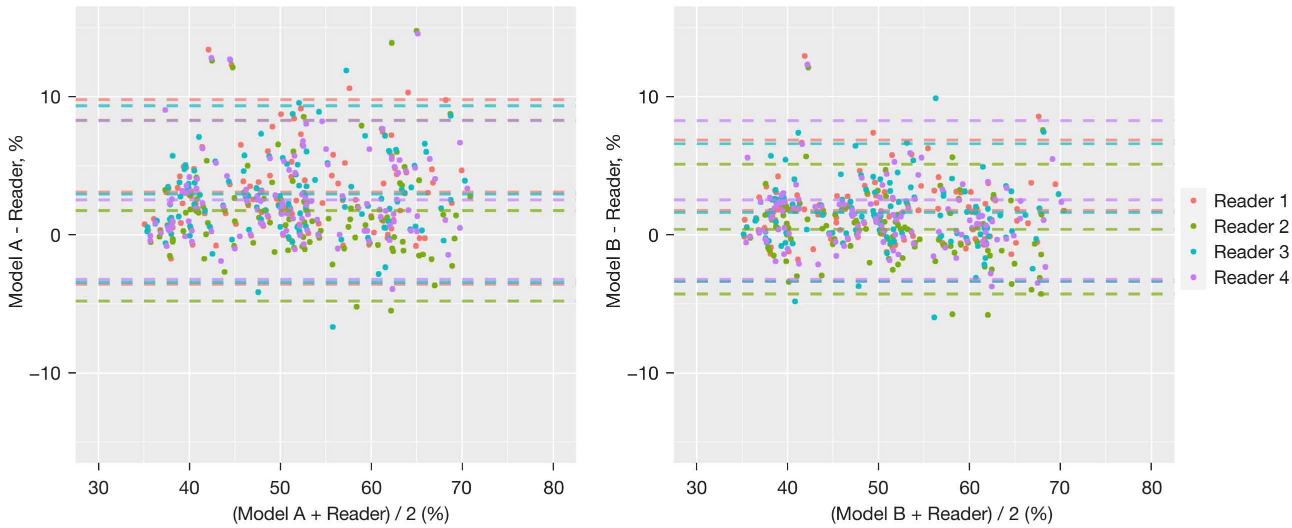


Figure 4 Bland-Altman plots of the two deep-learning models compared to manual measurements. The left panel shows the Bland-Altman plots of the initial model, and the references, and the right panel shows the plots of the advanced model and the references. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists.

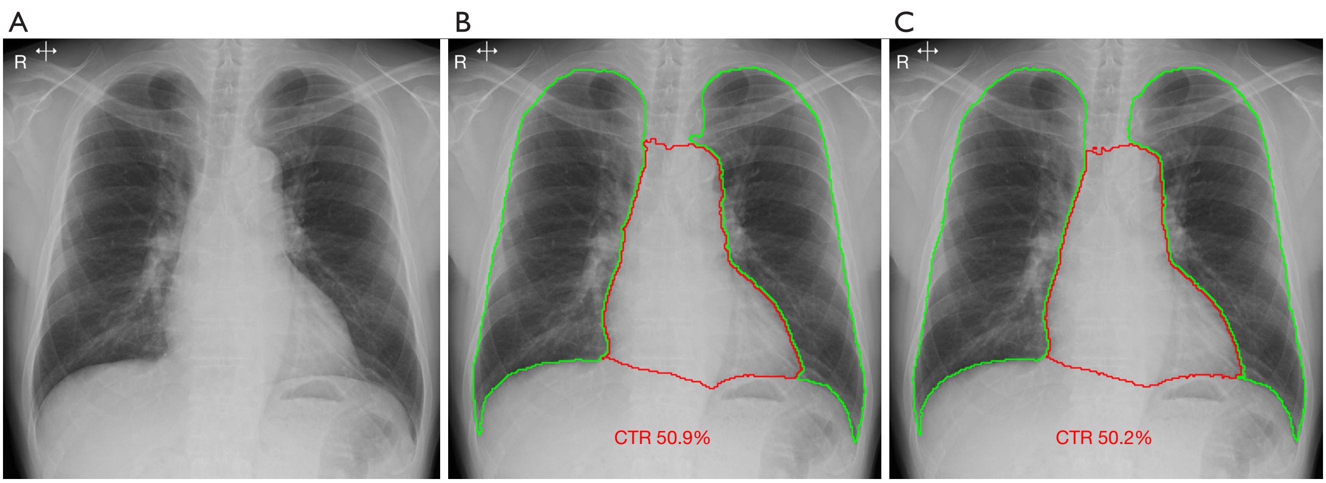


Figure 5 Example of chest radiograph showing good correlation with the references. (A) Original chest radiograph, (B) CTR estimation by the initial model, and (C) CTR estimation by the advanced model. The green line represents the contour of lung segmentation, and the red line represents the contour of heart segmentation. The mean CTR determined by four radiologists was 48.7% and the radiology report stated a CTR of 49%. CTR, cardiothoracic ratio.

studies have shown that measurement values may change with changes in patient conditions, highlighting the importance of the measurement of normal structures (11). In diagnostic imaging which focuses on detecting abnormalities, it can be a strain on the radiologist to simultaneously measure various normal structures. In addition, manual measurements performed by a radiologists can result in errors, including inter-reader or intra-reader error (12). Especially in longitudinal observations that investigate temporal changes in the same patient, inter-measurement discrepancies by various radiologists constitute a challenge that needs to be resolved. To reduce

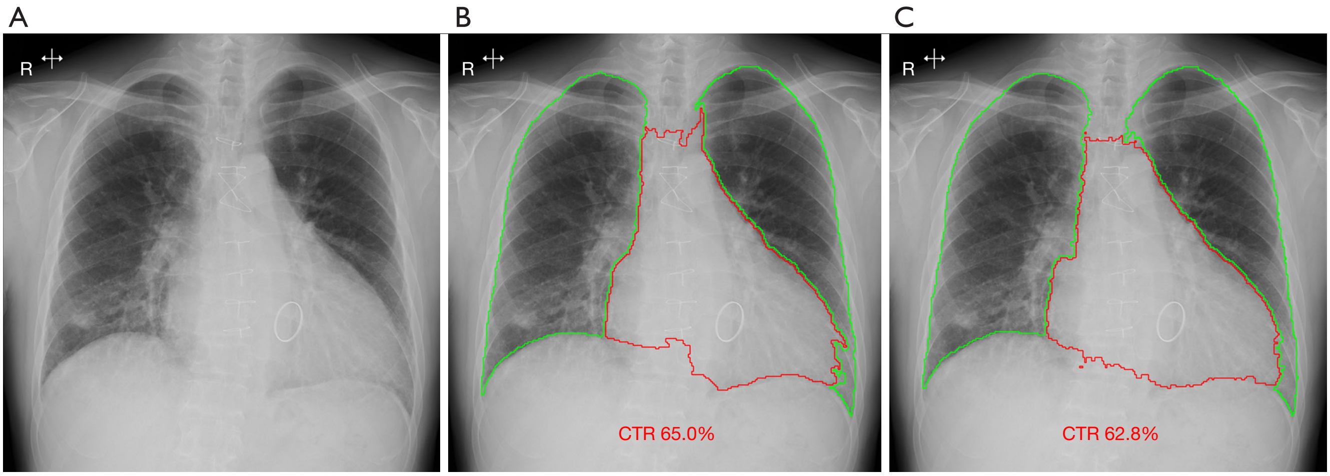


Figure 6 Example of chest radiograph showing poor correlation with the references. (A) Original chest radiograph, (B) CTR estimation by the initial model and (C) CTR estimation by the advanced model. The green line represents the contour of lung segmentation, and the red line represents the contour of heart segmentation. Compared to the initial model (B), the advanced model is more precise in the segment at the boundary between the left ventricle and the lung. The mean CTR determined by four radiologists was 58.1%, while the radiology report stated a CTR of 60.1%. CTR, cardiothoracic ratio.

workload and errors in measurement, the development of automatic measurement is necessary.

We attempted to construct a model for automatic measurement using U-net, a type of AI that has recently been shown to have high performance (13). To reduce the most effort-intensive segmentation work, we first built an initial model using an open-source training dataset and then constructed an advanced model by adding additional training data produced by the initial model to improve accuracy. While additional training can help the model adapt to the data that it is being trained on, using a limited number of institutions for training may result in overfitting to those specific institutions. However, even if limited to the institutions, this method has the potential to improve accuracy while minimizing the workload required for additional training.

In the two deep learning-based models for estimating the cardiothoracic ratio that were examined in this study, the results of the cardiothoracic ratio in the initial model tended to be slightly larger than that derived from manual measurement. This trend improved in the advanced model, which was trained with additional training data. The large cardiothoracic ratio in the initial model may be due to difficulties in accurately segmenting the boundary between the left ventricle or right atrium and the lung in the region close to the lung. On chest radiographs, faint increased density due to bilateral pericardial fat can often make it difficult to manually set the boundaries for cardiothoracic ratio measurement. However, the advanced model used in this study was able to make estimations that more closely resemble human manual measurements. The cardiothoracic ratio calculated by the models were slightly higher from that obtained through manual assessment. Consequently, when we used the common cut-off point of a cardiothoracic ratio ≥50% to define cardiomegaly, the results for the presence of cardiac enlargement may differ from those of manual measurements in certain cases.

Several methods for estimating the cardiothoracic ratio have been studied using various classical mathematical algorithm; however, in the current era of AI, automatic measurement by segments using deep learning with higher accuracy is expected to be applied to clinical practice (14-16).

This study has several limitations. First, it does not assess generalizability by studying data from other institutions, because it only attempts to improve accuracy by fitting with additional training data. It is possible that generalizability could be achieved by using this method to learn from data from multiple institutions. Second, it has not been determined whether results comparable to manual measurement can be obtained in a longitudinal study of the same patient. Third, patient clinical information, such as sex, age, and body shape, is different between the chest radiographs of the open-source dataset and the original, which may have affected the learning process.

Conclusions

The deep learning-based cardiothoracic ratio estimation model for chest radiographs showed good correlation with the manual measurements made by radiologists. The modified model, which used additional cases from the initial learning model, showed improved correlation with manual measurements compared to the initial model.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-23-187/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional board of Kanazawa University Hospital and individual consent for this retrospective analysis was waived.

References

1. Truszkiewicz K, Poręba R, Gać P. Radiological Cardiothoracic Ratio in Evidence-Based Medicine. J Clin Med 2021;10:2016.
2. Dimopoulos K, Giannakoulas G, Bendayan I, Liodakis E, Petraco R, Diller GP, Piepoli MF, Swan L, Mullen M, Best N, Poole-Wilson PA, Francis DP, Rubens MB, Gatzoulis MA. Cardiothoracic ratio from postero-anterior chest radiographs: a simple, reproducible and independent marker of disease severity and outcome in adults with congenital heart disease. Int J Cardiol 2013;166:453-7.
3. Bradley AJ, Carrington BM, Lawrance JA, Ryder WD, Radford JA. Assessment and significance of mediastinal bulk in Hodgkin's disease: comparison between computed tomography and chest radiography. J Clin Oncol 1999;17:2493-8.
4. Yotsueda R, Taniguchi M, Tanaka S, Eriguchi M, Fujisaki K, Torisu K, Masutani K, Hirakata H, Kitazono T, Tsuruya K. Cardiothoracic Ratio and All-Cause Mortality and Cardiovascular Disease Events in Hemodialysis Patients: The Q-Cohort Study. Am J Kidney Dis 2017;70:84-92.
5. Saiviroonporn P, Rodbangyang K, Tongdee T, Chaisangmongkon W, Yodprom P, Siriapisith T, Wonglaksanapimon S, Thiravit P. Cardiothoracic ratio measurement using artificial intelligence: observer and method validation studies. BMC Med Imaging 2021;21:95.
6. Hasan MA, Lee SL, Kim DH, Lim MK. Automatic evaluation of cardiac hypertrophy using cardiothoracic area ratio in chest radiograph images. Comput Methods Programs Biomed 2012;105:95-108.
7. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, Folio LR, Summers RM, Rubin DL, Lungren MP. Preparing Medical Imaging Data for Machine Learning. Radiology 2020;295:4-15.
8. Chen D, Liu S, Kingsbury P, Sohn S, Storlie CB, Habermann EB, Naessens JM, Larson DW, Liu H. Deep learning and alternative learning strategies for retrospective real-world clinical data. NPJ Digit Med 2019;2:43.
9. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kodera Y, Doi K. 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. AJR Am J Roentgenol 2000;174:71-4.
10. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015:234-41. Available online: https://arxiv.org/pdf/1505.04597.pdf
11. Raghu VK, Weiss J, Hoffmann U, Aerts HJWL, Lu MT. Deep Learning to Estimate Biological Age From Chest

Radiographs. JACC Cardiovasc Imaging 2021;14:2226-36.
12. Zhao B, Tan Y, Bell DJ, Marley SE, Guo P, Mann H, Scott ML, Schwartz LH, Ghiorghiu DC. Exploring intra- and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals. Eur J Radiol 2013;82:959-68.
13. Siddique N, Paheding S, Elkin CP, Devabhaktuni V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. IEEE Access 2021;9:82031-57.
14. Agrawal T, Choudhary P. Segmentation and classification on chest radiography: a systematic survey. Vis Comput 2023;39:875-913.
15. Lee WL, Chang K, Hsieh KS. Unsupervised segmentation of lung fields in chest radiographs using multiresolution fractal feature vector and deformable models. Med Biol Eng Comput 2016;54:1409-22.
16. Chamveha I, Promwiset T, Tongdee T, Saiviroonporn P, Chaisangmongkon W. Automated cardiothoracic ratio calculation and cardiomegaly detection using deep learning approach. Arxiv 2020:1-11. Available online: https://arxiv.org/pdf/2002.07468.pdf

Cite this article as: Yoshida K, Takamatsu A, Matsubara T, Kitagawa T, Toshima F, Tanaka R, Gabata T. Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement without self-annotation. Quant Imaging Med Surg 2023;13(10):6546-6554. doi: 10.21037/qims-23-187

Supplementary

Table S1 Parameters of U-net model

Parameters	Value
Learning rate	0.0001
Minibatch size	64
Epochs	1,000
Optimizer	Adam
Loss function	Cross entropy loss

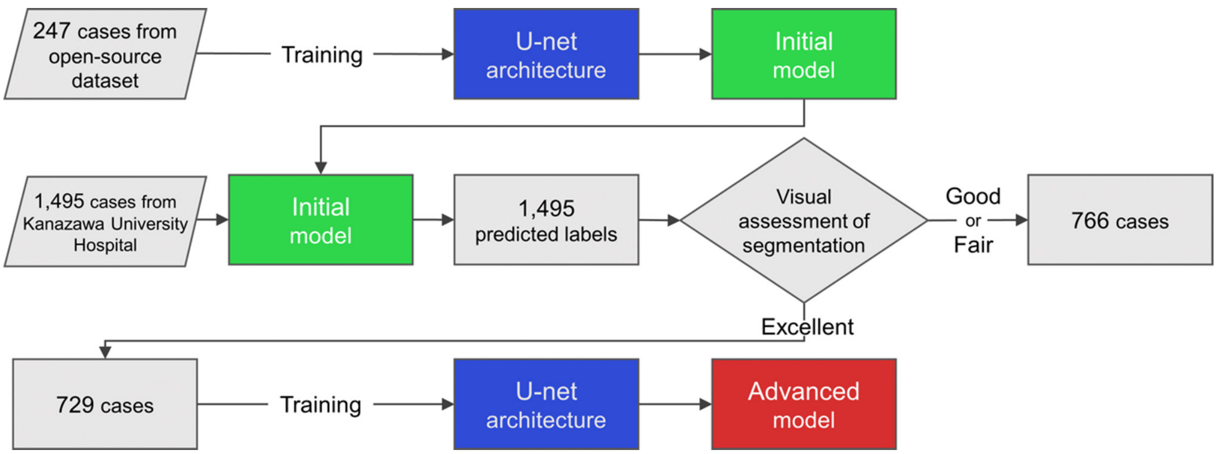
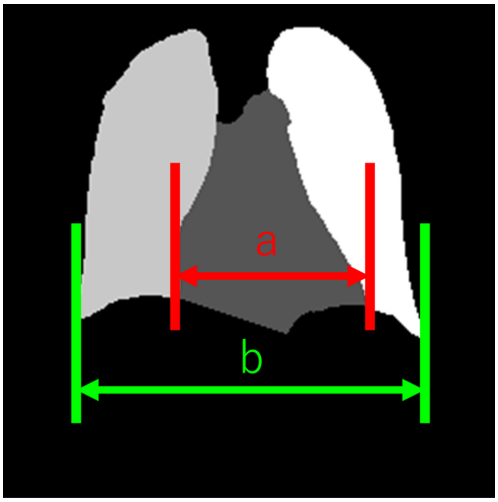


Figure S1 Flowchart of model development (initial model and advanced model).



Cardiothoracic ratio = a/b

Figure S2 Output of cardiothoracic-ratio measurement. The transverse axis length of the heart was represented as “a” and that of the thorax was represented as “b”.

 https://dx.doi.org/10.21037/qims-23-187

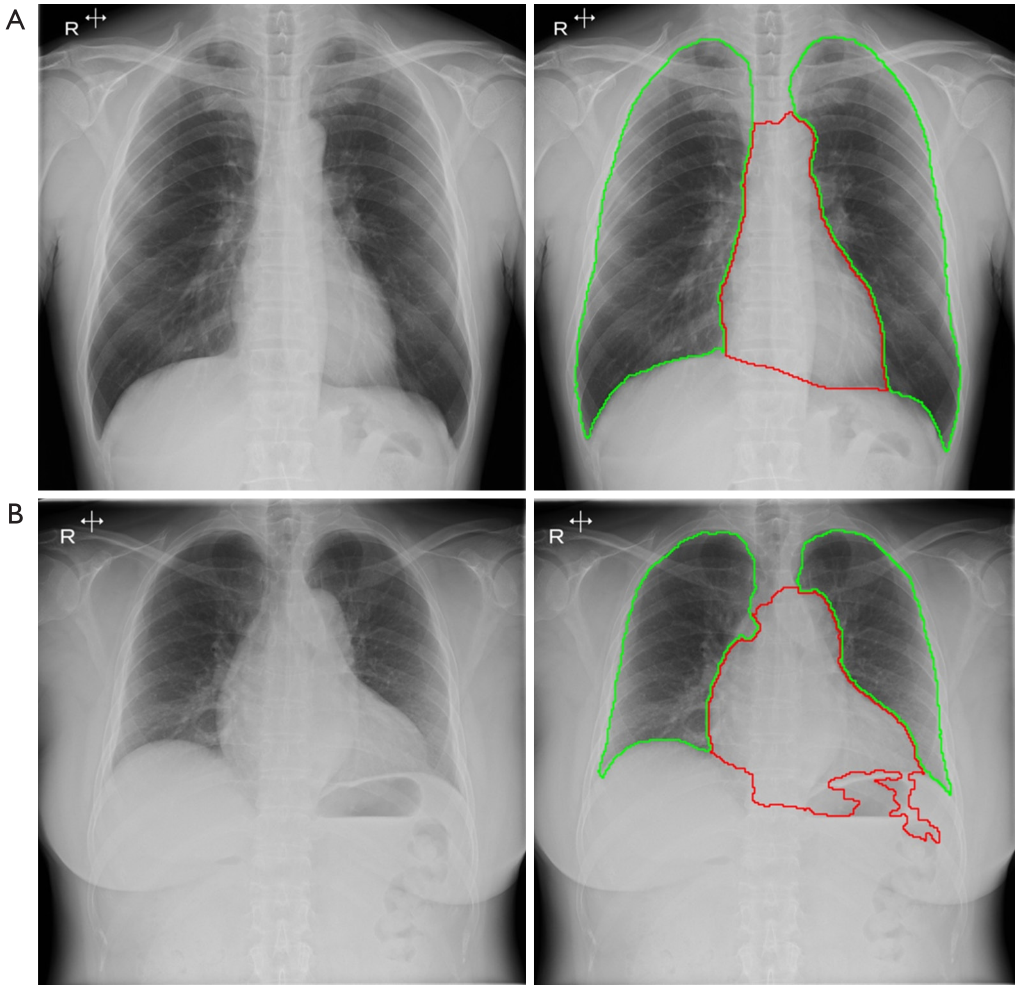


Figure S3 Examples of excellent and fair segmentation of chest radiographs performed by initial mode. (A) An excellent segmentation case. On the left is the original chest radiograph, and on the right is the image with contour information for the segmentation of lungs (green line) and heart (red line). (B) A fair segmentation case. On the left is the original chest radiograph, and on the right is the image with contour information for the segmentation of lungs (green line) and heart (red line). There are irregularities in the lower margin of the heart segment and incorrect segmentations where the gastric bubbles overlap.

https://dx.doi.org/10.21037/qims-23-187