IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 17, 2019, accepted February 6, 2019, date of current version April 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2900053

Automatic Cardiothoracic Ratio Calculation With Deep Learning

ZHENNAN LI[1], ZHIHUI HOU[1], CHEN CHEN[2], ZHI HAO[2], YUNQIANG AN[1], SEN LIANG[2], AND BIN LU[1]

[1]Department of Radiology, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, 100037, China
[2]Infervision, Beijing, China

Corresponding author: Bin Lu (blu@vip.sina.com)

This paper was supported in part by the Ministry of Science and Technology of China, in part by the National Key Research and Development Project under Grant 2016YFC1300400, and in part by CAMS Innovation Fund under Grant 2016-I2M-1-011.

ABSTRACT Deep learning is a growing trend in medical image analysis. There are limited data of deep learning techniques applied in Chest X-rays. This paper proposed a deep learning algorithm for cardiothoracic ratio (CTR) calculation in chest X-rays. A fully convolutional neural network was employed to segment chest X-ray images and calculate CTR. CTR values derived from the deep learning model were compared with the reference standard using Bland–Altman analysis and linear correlation graphs, and intra-class correlation (ICC) analyses. Diagnostic performance of the model for the detection of heart enlargement was assessed and compared with other deep learning methods and radiologists. CTR values derived from the deep learning method showed excellent agreement with the reference standard, with mean difference 0.0004 ± 0.0133, 95% limits of agreement -0.0256 to 0.0264. Correlation coefficient between deep learning and reference standard was 0.965 ($P < 0.001$), and ICC coefficient was 0.982 (95% CI 0.978–0.985) ($P < 0.001$). Measurement time by deep learning was significantly less than that of the manual method [0.69 (0.69–0.70) VS 25.26 (23.49–27.44) seconds, $P < 0.001$]. Diagnostic accuracy, specificity, and positive predictive value were comparable between the two methods. However, deep learning showed relatively higher sensitivity and negative predictive value (97.2% vs 91.4%, $P = 0.004$; and 96.0% vs 89.0%, $P = 0.006$; respectively) compared with the manual method. Performance of this computer-aided technique was demonstrated to be more reliable, time and labor saving than that of the manual method in CTR calculation.

INDEX TERMS Machine learning, deep learning, object segmentation, X-rays application, cardiomegaly.

I. INTRODUCTION

Chest X-ray is the most commonly used modality in clinical practice for screening for lung and heart diseases. As a first-line imaging tool, Chest X-ray is cheap, readily available and easy to interpret. Cardiothoracic Ratio (CTR) derived from Chest X-ray is a widely used radiographic index to assess cardiac size and provide prognostic information in both congenital and acquired heart diseases [1]–[3]. Enlarged CTR often indicates structural cardiac abnormalities and cardiac enlargement related diseases, and has been shown to associate with functional status and adverse clinical outcomes [3]. However, manual calculation of CTR in routine clinical practice is subjective and time consuming, and may introduce large variations across interpreters. Emerging computerized tools applied to medical image processing help improve diagnosis and simplify expert workflow. As for Chest X-ray, however, accurate interpretation in computer-aided diagnostic schemes is extremely challenging due to the complexity of anatomical structures. Lung and heart boundaries need to be accurately identified in Chest X-rays before further analysis including CTR calculation. Automatic lung and heart fields segmentation in computer-aided diagnostic systems is an important and challenging task.

Recently, with technical advances in computer science, machine learning has been increasingly used in the medical imaging field. As one type of machine learning, deep learning is emerging as the leading machine-learning algorithm in medical imaging analysis and computer vision domains and is a growing trend in medical big data analysis [4]–[6]. It has been successfully applied in medical imaging field with impressive performance, including image classification,

The associate editor coordinating the review of this manuscript and approving it for publication was Tomasz Trzcinski.

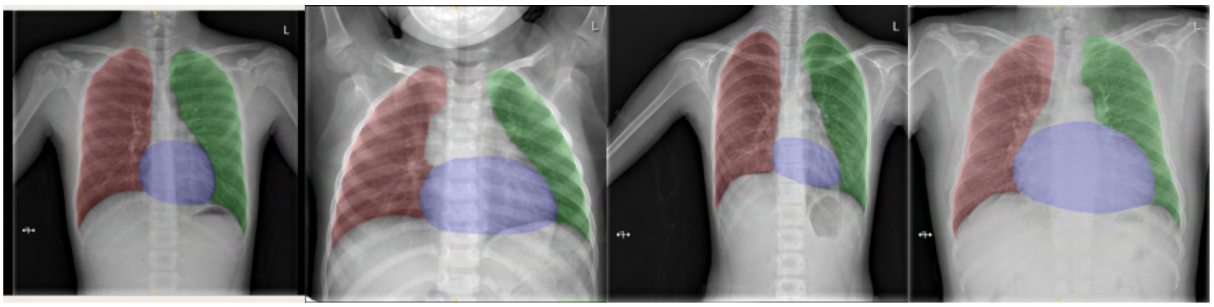


FIGURE 1. Examples of various heart and lung segmentation by radiologists. These figures show high shape variability and boundary ambiguity of the lung field and heart borders due to individual differences.

segmentation and object detection [7]–[12]. A typical approach of deep learning called convolutional neural networks (CNNs) automatically learns mid-level and high-level abstractions obtained from raw data, and has demonstrated to be a powerful method for tasks related to medical image processing [5]–[6]. Automated segmentation of liver and tumors and other structures in the brain from CT and MR imaging has been reported using CNNs with promising results [9]–[11]. However, there has been limited application of the exciting deep learning technique in Chest X-rays [13]–[14]. Automated analysis of Chest X-ray images can be well performed by the deep learning method. Reference [25] attempts to use the deep learning classic structure U-net on a limited data set to calculate CTR, and obtains good results.

In the present study, we proposed a deep learning algorithm to segment lung and heart regions in Chest X-rays and calculate CTR accordingly. The performance of this computerized method was evaluated in terms of accuracy and speed of computation. This work is a demonstrations of deep learning applied to cardiovascular problem diagnosis on X-ray images.

II. MATERIALS AND METHODS

A. DATASET

This study was performed at the Radiology Imaging Center of our hospital. The study complied with the Declaration of Helsinki and was approved by the Institutional Review Board. A total of 5000 postero-anterior chest X-rays of 5000 patients in different age groups were consecutively collected from picture archiving and communication systems. Substantial varieties were observed across different patient age groups, pathological conditions, as well as imaging sizes in this dataset. In order to protect the patients' privacy (e.g., patient name, age, date of examination), all the images were desensitized before using.

B. TRAINING AND VALIDATION SET

This dataset was randomly split into two parts: 4000 (80%) for training and 1000 (20%) for validation. The training set was used to train our model while the validation set was used to estimate how well our model had been trained. In this work, the images for training and validation were labeled with lung and heart segmentation. As shown in Figure 1, the regions of red, green, blue are represented for the right lung, left lung and heart respectively. In this process, a software called ITK-SNAP [15] was adopted for experts to draw and edit the masks conveniently.

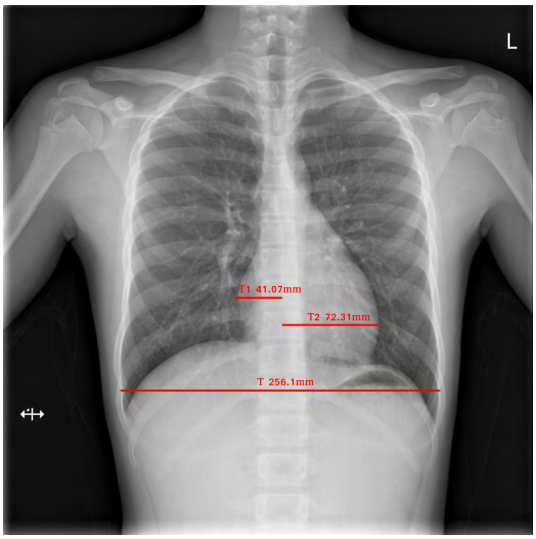


FIGURE 2. Measurement of transverse diameter of thorax and heart. The longest line measures the internal diameter of the chest as "T". Transverse heart diameter is determined by summing the lengths of "T1" and "T2" where "T1" is the transverse diameter on the right side of the heart and "T2" is the one on the left side.

C. TESTING SET

Additional 500 X-ray images were later collected to assess the CTR calculation performance of our model, for which, CTR calculation both by deep learning and manual method was performed. We invited two experienced radiologists to measure the transverse thoracic diameter and the transverse heart diameter independently. An example of measurement is shown in Figure 2. Reference standard of each CTR was determined by calculating the average CTR from the two radiologists' measurements. A third measurement was performed to achieve a final consensus result if the difference of CTR values was less than 0.05.In addition, a separate radiologist was also invited to measure the 500 X-ray images and calculate CTR independently, performance of this manual method was compared with the deep learning model. Furthermore, 100 chest X-ray images were randomly selected to evaluate intra- and inter-observer agreement of manual measurements of CTR.

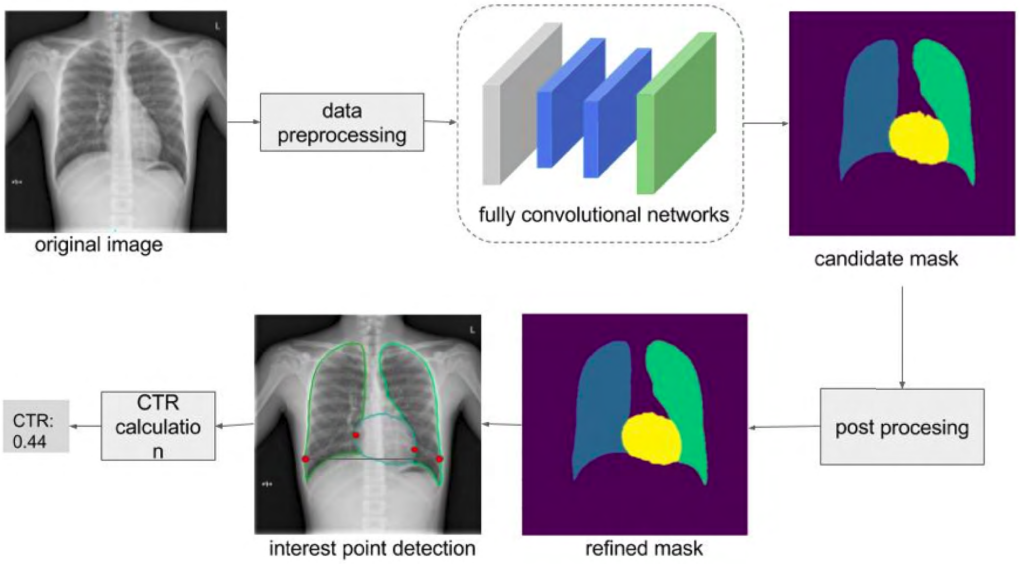


FIGURE 3. Overview of our proposed framework.

D. DEEP LEARNING

Our automatic calculation of CTR was based on accurate lung and heart segmentation. It mainly contained two steps: 1) lung and heart segmentation, 2) CTR calculation. In this work, we used deep learning techniques to automatically perform segmentation task. Specifically, a fully convolutional neural network inspired from Unet [16] was employed to segment left lung, right lung and heart region from an X-ray image. After extracting contours from left-right lung and heart segmentation masks, interest points were located to measure heart transverse diameters and chest transverse diameters for CTR calculation. Figure 3 shows overview of our proposed framework.

E. DATA PREPROCESSING

Due to heterogeneity observed in the scale of image intensity, we performed gray-scale transformation for image normalization. For every single image, we used window center wc and window width ww stored in its dicom lookup table to map the output pixel values to a gray scale of $0-255$. For each pixel p in a single dicom, the gray transformation process to p was defined as follows:

1) $thresh_{min} = wc - ww/2$ and $thresh_{max} = wc + ww/2$
2) $p = 255$, if $p > thresh_{max}$
3) $p = 0$, if $p < thresh_{min}$
4) $p = \frac{(p-thresh_{min})*255}{thresh_{max}-thresh_{min}}$, if $thresh_{min} < p < thresh_{max}$

where wc is the window center and and ww denote its window width of the dicom image. As a result, this process adjusted all the images to achieve the desired effect for visualizing heart and lung field.

Since chest X-ray images we collected exhibit substantial variety in image sizes where image widths and heights ranging from 2000 to 3100 pixels, image resize process was required to unify the image size before feeding images into the network. In order to keep sufficient image visual details for delineating heart and lung contours, we simply cropped and rescaled images. Firstly, we center cropped to the largest square region of images. Secondly, we downscaled all images to 800×800 pixels.

While data augmentation has been widely used in deep learning to improve segmentation performance, we found that augmenting images by rotating and zooming images did not improve our segmentation results in our early trials. This may be due to the fact that images in the dataset we observed are rarely rotated. Hence, data augmentation was not considered.

F. FULLY CONVOLUTIONAL NEURAL NETWORK FOR SEGMENTATION

We adopted a 2D U-Net [16] architecture to perform lung segmentation and heart segmentation in our work. Our customized U-net (shown in Figure 4) model is a 32-layer fully convolutional neural network. It takes a chest X-ray image as an input and outputs the probabilities of being each class (background, left lung, right lung, and heart) at each pixel. Unlike vanilla U-Net, we replaced the transposed deconvolution network with upsampling and convolution layers following the strategy proposed in [17]. Subsequent experiments proved that these changes could improve the network compared to vanilla U-net. The main elements of architecture are depicted in Figure 4. All convolutional layers use batch normalization [18], ReLUs [19] and zero-padding.

The whole network consists of two parts: one is feature extraction branch and the other is feature aggregation branch. In the left part of the network architecture, images are fed through a series of convolution and pooling layers where it can extract features at different scales and different levels, from low-level features (e.g. edge, corner etc.) to high-level semantic information such as shape. On the right, the network adopts multi-scale feature aggregation fashion to nonlinearly combine feature learned at different image levels via blocks of upsampling, concatenating and convolutional layers. Through this ''skip connection'' procedure,

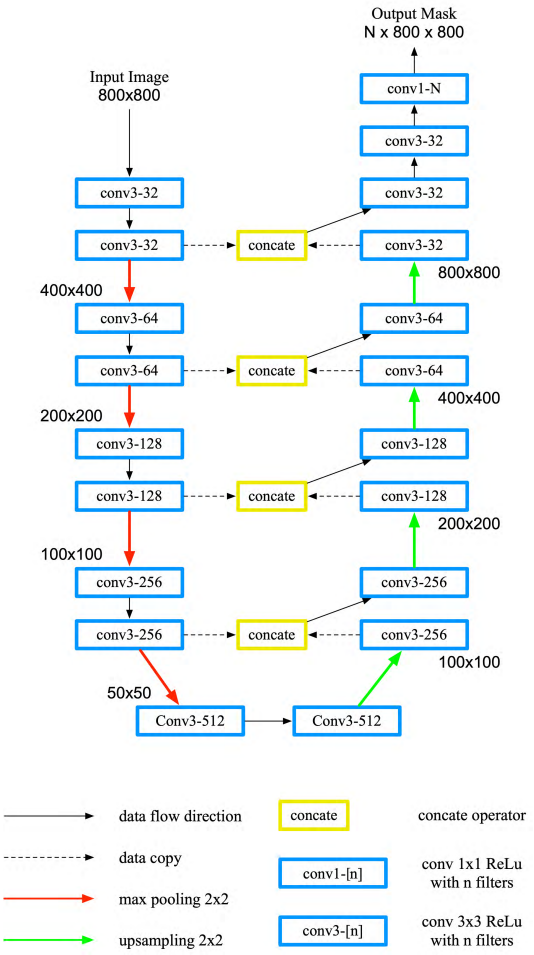


FIGURE 4. Our adapted U-net architecture. Each blue box corresponds to a single convolutional layer. The size and number of convolution kernels are stated inside each box(e.g. conv3-64 means there are 64 3x3 convolutional filters). Each yellow box represents concatenate operation which simply concatenates a copy from left branch output feature maps with upscaled feature maps from the right branch. Red Arrows denote max-pooling layers while blue arrows correspond to upsampling layers. Both max-pooling and upsampling will change the output feature map sizes. Sizes of feature maps from different level are only provided when it changed.

spatial information loss caused by max-pooling layers will be recovered level by level. In the end, the output prediction for each class carries class-specific shape information with fine details, which benefits for precise segmentation.

Given a chest X-ray image from a patient, our trained network will output the probability of each pixel for each class via a softmax function. The final segmentation made by the model is performed by assigning each pixel to the class with the highest probability. The borders of segmentation masks are then refined by dense Conditional Random Field(dense CRF). The details of how dense CRF works is illustrated in the next section.

G. POST-PROCESSING

In the post-processing stage, we use fully-connected CRF [23] to smooth region boundaries of segmentation masks. As a result, any holes occurred in the candidate masks are filled and mask boundaries will become more smooth. After that, the largest region is kept to mask organs (left lung field, right lung field, and heart). Fully-connected CRF calculates pairwise energy between all pairs of pixels in the image in order to refine dense segmentation masks. The energy function is described as follows:

$$E(x) = \sum_i \theta_u(x_i) + \sum_{i,j} \theta_p(x_i, x_j). \tag{1}$$

where x is the pixel-wise label assignment. We use the label assignment probability $p(x_i)$ to calculate the unary potential $\theta_u(x_i)$:

$$\theta_u(x_i) = -logp(x_i). \tag{2}$$

where $p(x_i)$ is actually the final output from our U-net segmentation network. The second term of this equation is pairwise potential which allows for efficient inference even in a fully-connected graph. The second term formula is as follows:

$$\theta_p(x_i, x_j) = \left[w_1 exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_\alpha^2} - \frac{\|I_i - I_j\|^2}{2\alpha_\beta^2} \right) + w_2 exp\left(-\frac{\|p_i - p_j\|^2}{2\alpha_\gamma^2} \right) \right]_{x_i \neq x_j}. \tag{3}$$

where p denotes pixel position and I denotes intensity. This formula consists of two penalty terms weighted by w_1 and w_2, which only taking those nodes with different labels into account. The first one models the similarity (both pixel location and RGB color/intensity and spatial proximity) while the second one only cares about spatial proximity. The degree of the similarity is controlled by the scale of Gaussian Kernel σ_α, σ_β, and σ_γ. By minimizing the energy, fully-connected CRFs force pixels with similar color/intensity and position to have similar labels and enforces smoothness.

H. CTR CALCULATION

Cardiothoracic ratio was automatically computed based on these boundaries derived from organ masks. We identified the internal diameter of the chest as "T" by finding the longest horizontal distance between right lung and left lung. Instead of adding the right side of the heart transverse diameter "T1", and left side of heart transverse diameter "T2", we simply get the transverse diameter of heart by measuring the horizontal distance "Th" between the most left points and the most right points on the heart contours(see Figure 5). CTR was calculated as the ratio of the transverse diameter of the heart to the internal diameter of a chest. The formula is shown as follows:

$$CTR = Th/T. \tag{4}$$

I. EXPERIMENTAL SETTING

The weights of the network were initialized using Xavier Initialization method [20]. The whole network was trained end-to-end using batch stochastic gradient descent (SGD) [21]

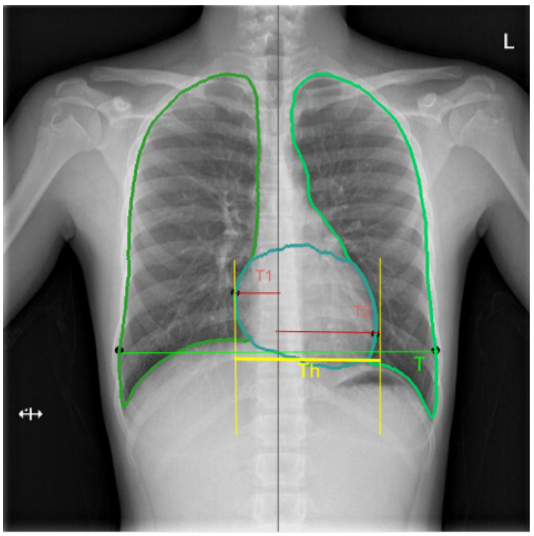


FIGURE 5. CTR calculation.

method with a batch size of 4 and we set the initial learning rate to be 0.001. We trained the model with cross-entropy loss. To accelerate the training process, we employed batch normalization [18]. Our proposed model was implemented in MXNet [22] with two 1080 Ti GPUs. Our trained network output the probability of each pixel for each class via a softmax function. The final segmentation made by the model is performed by assigning each pixel to the class with the highest probability. The borders of segmentation masks would be refined by dense conditional random fields (CRF) [23].

J. STATISTICAL ANALYSIS

Statistical analysis was performed with SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). Kolmogorov Smirnov tests were used to evaluate the distribution of the data. Paired t test (for normally distributed variables) and Wilcoxon signed ranks test (for not normally distributed variables) were used to compare CTR values as well as measurement time between paired groups. Reliability of CTR calculation was determined using intra-class correlation (ICC) analyses. Excellent agreement was defined as an ICC coefficient of > 0.8.

In addition, diagnostic performance of the deep learning model for detection of cardiomegaly was assessed using a binary approach (using CTR > 0.5 as a cut-off), and was compared with that of manual method.

III. RESULTS

A. CTR DERIVED FROM OUR METHOD AND OTHER DEEP LEARNING METHODS

In the testing set, all 500 chest X-rays were successfully processed by the deep learning models, including automatic lung and heart segmentation and CTR calculation. Since [25] directly uses the U-net architecture, our comparison with the U-net is equivalent to the comparison with [25]. Table 1 shows the comparison results of our proposed method with U-net and U-net + CRF. Our method shows better performance in all aspects than U-net.

B. CTR DERIVED FROM THE DEEP LEARNING MODEL AND MANUAL METHOD

CTR values derived from the deep learning method showed excellent agreement with the reference standard (as shown in Table 2 and Figure 5). Paired difference between deep learning and reference standard showed no statistical significance ($P = 0.552$), with mean difference 0.0004 ± 0.0133, 95% limits of agreement -0.0256 to 0.0264. Correlation coefficient (r^2) between deep learning and reference standard was 0.965 ($P < 0.001$), and intra-class correlation (ICC) coefficient was 0.982 (95% CI 0.978-0.985) ($P < 0.001$). As for manual method, however, mean difference was -0.0083 ± 0.0187 with statistical significance ($P < 0.001$), 95% limits of agreement -0.0450 to 0.0284. Correlation coefficient (r^2) between manual method and reference standard was 0.929 ($P < 0.001$), and ICC coefficient was 0.957 (95% CI 0.928-0.972) ($P < 0.001$).

C. INTRA- AND INTEROBSERVER REPRODUCIBILITY OF MANUAL MEASUREMENTS OF CTR

Intra- and interobserver reproducibility of manual measurements of CTR was assessed in 100 randomly selected chest

TABLE 1. Diagnostic performance between our method and other deep learning methods.

	U-net	U-net + CRF	Our method
Accuracy	94.5%	94.9%	95.3%
Sensitivity	96.1%	96.1%	97.2%
Specificity	92.2%	93.2%	92.7%
PPV	94.4%	95.1%	94.8%
NPV	94.5%	94.6%	96.0%

PPV = positive predictive value; NPV = negative predictive value.

TABLE 2. Diagnostic performance of deep learning and manual method for detection of heart enlargement.

	Deep learning method, % (n/N)	Manual method, % (n/N)	P value
Accuracy	95.3 (477/500)	93.8 (469/500)	0.259
Sensitivity	97.2 (284/292)	91.4 (267/292)	0.004
Specificity	92.7 (193/208)	97.1 (202/208)	0.107
PPV	94.8 (284/299)	97.8 (267/273)	0.116
NPV	96.0 (193/201)	89.0 (202/227)	0.006

PPV = positive predictive value; NPV = negative predictive value. Deep learning method refers to our method proposed in this paper.

TABLE 3. Paired t test results and agreement between paired groups.

	Mean values	Paired differences	P values	95% limits of agreement	ICC coefficient (95% CI)
Deep learning	0.5233 ± 0.0711	0.0004 ± 0.0133	0.552	-0.0256 ∼ 0.0264	0.982 (0.978 ∼ 0.985)
Manual method	0.5146 ± 0.0700	-0.0083 ± 0.0187	< 0.001	-0.0450 ∼ 0.0284	0.957 (0.928 ∼0.972)
Intra-observer(n=100)	0.5152 ± 0.0674	-0.0024 ± 0.0141	0.099	-0.0299 ∼ 0.0253	0.978 (0.968 ∼ 0.985)
Inter-observer(n=100)	0.5282 ± 0.0647	0.0163 ± 0.0196	< 0.001	-0.0221 ∼ 0.0547	0.927 (0.704 ∼ 0.970)

Values are reported as mean ± standard deviation. ICC = intra-class correlation. 95%CI = 95% confidence interval.

X-rays in our study. As shown in Table 3 and Figure 5, good intra-observer reproducibility was observed for manual measurement of CTR. Mean difference between twice measurements by one observer was -0.0024 ± 0.0141 with no statistical significance ($P = 0.099$), 95% limits of agreement -0.0299 to 0.0253. Correlation coefficient (r^2) between the intra-observer measurements was 0.957 ($P < 0.001$), and ICC coefficient was 0.978 (95% CI 0.968-0.985) ($P < 0.001$). As for interobserver reproducibility, however, mean difference between the two observers was 0.0163 ± 0.0196 with statistical significance ($P < 0.001$), 95% limits of agreement -0.0221 to 0.0547. Correlation coefficient (r^2) between the 2 observers was 0.913 ($P < 0.001$), and ICC coefficient was 0.927 (95% CI 0.704-0.970) ($P < 0.001$).

In addition, measurement times of CTR both by deep learning and manual method were recorded and compared in the 100 chest X-rays. Measurement time by deep learning method was less than 1 second for all the samples, which was significantly less than that of manual method [0.69 (0.69 - 0.70) VS 25.26 (23.49 - 27.44) seconds, $P < 0.001$].

D. DIAGNOSTIC PERFORMANCE

Diagnostic performance of the deep learning method for detection of cardiomegaly (using CTR > 0.5 as a cut-off) was assessed and compared with those of manual method. In this analysis, CTR values derived from the reference standard were taken as reference. There was no statistically significant difference with regard to diagnostic accuracy, specificity and positive predictive value between the two methods (Table 2). However, deep learning method showed relative higher sensitivity and negative predictive value (97.2% vs 91.4%, $P = 0.004$; 96.0% vs 89.0%, $P = 0.006$; respectively) compared with manual method.

IV. DISCUSSION

This is the first study to present a deep learning technique to segment lung and heart regions in Chest X-rays and calculate CTR automatically. Deep learning is a subfield of machine learning characterized by the use of multilayer artificial neural networks, originally inspired by human neural systems. A classical deep learning architecture known as the convolutional neural networks (CNNs) has recently proved to be powerful tools for a broad range of computer vision tasks. With availability of big medical data, enhanced computing power, and new algorithms to train the CNNs, CNNs are increasingly applied to medical image processing with impressive performances, such as detection of mammographic lesions [7], lung nodule segmentation from computed tomography (CT) images [8], automatic segmentation of magnetic resonance brain images [10], [11], classification of interstitial lung diseases patterns with high-resolution chest CT [12]. To date, however, there has been limited research of CNNs applied in Chest X-rays. In this study, we proposed a fully convolutional neural network to segment chest X-ray images and calculate CTR automatically.

In this study, we used large training set to train and validate the deep learning model, all the images for training and validation were labeled with lung and heart segmentation. Our adapted U-net architecture exhibited excellent performance in chest X-ray image segmentation. Based on precise segmentation, CTR calculation was performed automatically. In the test set, we demonstrated excellent performance of the deep learning model, CTR values derived from the deep learning method showed excellent agreement with the reference standard, the intra-class correlation coefficient achieved 0.982 between the deep learning method and reference standard. As for manual method, however, slight mean difference was observed with statistical significance between manual method and reference standard, even though their agreement was high.

In addition, measurement time by deep learning method was less than one second for all the CTR calculation, suggesting that the computerized approach is more reliable, time and labor saving compared with traditional manual method.

This study demonstrated excellent diagnostic performance for both deep learning and manual method for detection of heart enlargement, diagnostic accuracy, specificity as well as positive predictive value were comparable between the two methods. It is interesting to note that, deep learning method showed relative higher sensitivity and negative predictive value compared with manual method in this analysis. Mean CTR values derived from the manual method were slightly less than that of reference standard with statistical significance. This might lead to relatively lower diagnostic sensitivity and negative predictive value for detection of heart enlargement. And variations across interpreters should be taken into account.

V. STUDY LIMITATIONS

There were several limitations in the present study. First, the deep learning model was designed for automatic segmentation of heart and lung in chest X-rays and CTR calculation, other image interpretations were beyond the scope of this article. Whether the deep learning technique is well suited to identify various pathologic findings in chest X-rays merits further investigation. Second, having a balanced dataset of

categories is important for the network [13], [24]. Although diverse pathological conditions were included in our training set, there were not enough examples to train for some rare pathology, such as images with obvious fat pad of pericardium and pleural effussion. The model sometimes has limited ability to distinguish clearly fat pad of pericardium from the real left heart boundary, and lung region can be wrongly identified when there exists obvious pleural effussion. CTR calculated by the model in these conditions might be inaccurate, and might generate more false positives than human readers in detection of cardiac enlargement. Additional training cases from these relative rare conditions might help form a robust training set for the network to learn from. Third, performance of the deep learning model for CTR calculation was demonstrated excellent in this retrospective study, whether the model can be well applied in clinical practice should be prospectively validated in further study. In addition, other deep neural network architectures can also be developed and compared with the present model.

VI. CONCLUSIONS

The deep learning model was successfully employed to segment chest X-ray images and calculate CTR automatically. Performance of this computer-aided technique was demonstrated to be more reliable, time and labor saving compared with the manual method in routine clinical practice. Deep learning techniques might be powerful tools applied in cardiovascular problem diagnosis.

ACKNOWLEDGMENT

The authors are grateful to the Infervision Corporation for their help with the deep learning algorithm used in this research. *(Zhennan Li and Zhihui Hou contribute equally to this work.)*

REFERENCES

[1] M. T. Kearney *et al.*, "Predicting death due to progressive heart failure in patients with mild-to-moderate chronic heart failure," *J. Amer. College Cardiol.*, vol. 40, no. 10, pp. 1801–1808, 2002.

[2] F. A. Hubbell, S. Greenfield, J. L. Tyler, K. Chetty, and F. A. Wyle, "The impact of routine admission chest X-ray films on patient care," *New England J. Med.*, vol. 312, pp. 209–213, Jan. 1985.

[3] K. Dimopoulos *et al.*, "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.*, vol. 166, pp. 453–457, Jun. 2013.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[5] G. Chartrand *et al.*, "Deep learning: A primer for radiologists," *Radiographics*, vol. 37, no. 7, pp. 2113–2131, 2017.

[6] J. G. Lee *et al.*, "Deep learning in medical imaging: General overview," *Korean J. Radiol.*, vol. 18, no. 4, pp. 570–584, 2017.

[7] T. Kooi *et al.*, "Large scale deep learning for computer aided detection of mammographic lesions," *Med. Image Anal.*, vol. 35, pp. 303–312, Jan. 2017.

[8] S. Wang *et al.*, "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Med. Image Anal.*, vol. 40, pp. 172–183, Aug. 2017.

[9] M. Drozdzal *et al.*, "Learning normalized inputs for iterative estimation in medical image segmentation," *Med. Image Anal.*, vol. 44, pp. 1–13, Feb. 2018.

[10] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016.

[11] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.

[12] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1207–1216, May 2016.

[13] M. Cicero *et al.*, "Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs," *Invest. Radiol.*, vol. 52, no. 5, pp. 281–287, 2017.

[14] A. Rajkomar, S. Lingam, A. G. Taylor, M. Blum, and J. Mongan, "High-throughput classification of radiographs using deep convolutional neural networks," *J. Digit. Imag.*, vol. 30, no. 1, pp. 95–101, 2017.

[15] P. A. Yushkevich *et al.*, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage*, vol. 31, pp. 1116–1128, Jul. 2006.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Cham, Switzerland: Springer, 2015, pp. 234–241.

[17] S. Honari, J. Yosinski, P. Vincent, and C. Pal, "Recombinator networks: Learning coarse-to-fine feature aggregation," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 5743–5752.

[18] S. Ioffe and C. Szegedy. (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift." [Online]. Available: https://arxiv.org/abs/1502.03167

[19] A. Krizhevsky and I. G. E. Sutskever Hinton, "Image Net classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*. Red Hook, NY, USA: Curran Associates, Inc., 2012, pp. 1097–1105.

[20] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249–256, May 2010.

[21] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," in *Neurocomputing: Foundations of Research*, Cambridge, MA, USA: MIT Press, 1988, pp. 533–536.

[22] T. Chen *et al.* (2015). "MXNet: A flexible and efficient machine learning library for heterogeneous distributed systems." [Online]. Available: https://arxiv.org/abs/1512.01274

[23] P. Krähenbühl and V. Koltun, "Efficient inference in fully connected CRFs with Gaussian edge potentials," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 109–117.

[24] C. Huang, Y. Li, C. C. Loy, and X. Tang, "Learning deep representation for imbalanced classification," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 5375–5384.

[25] Q. Que *et al.*, "CardioXNet: Automated detection for cardiomegaly based on deep learning," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Honolulu, HI, USA, Jul. 2018, pp. 612–661. doi: 10.1109/EMBC.2018.8512374.

ZHENNAN LI was born in Xinxiang, Henan, China, in 1988. He received the B.S.Med. degree in clinical medicine from the Chengdu University of Traditional Chinese Medicine, Chengdu, China, in 2010, and the M.D. degree in clinical medicine from Chinese Academy of Medical Science, Peking Union Medical College, Beijing, China, in 2016.

From 2017 to 2019, he was a Postdoctoral Fellow with the Department of Radiology, Fuwai Hospital, China. His research interest includes the development of coronary atherosclerosis, effects of statin therapy on progression of mild noncalcified coronary plaque assessed by serial coronary computed tomography angiography, predictive models for aortic intramural hematoma, large scale multicenter prospective cohort study in cardiovascular diseases, and machine learning approaches in cardiovascular imaging.

Dr. Li was a recipient of the China National Scholarship, in 2012.

ZHIHUI HOU was born in Kaifeng, Henan, China, in 1984. He received the B.S.Med. degree in clinical medicine from Wuhan University, Wuhan, China, in 2008, and the M.D. degree in clinical medicine from Chinese Academy of Medical Science, Peking Union Medical College, Beijing, China, in 2013. He is currently an Attending Physician with the Department of Radiology, Fuwai Hospital, China. His research interests include prognostic value of coronary CT angiography and calcium score for major adverse cardiac events in outpatients, large scale multicenter prospective cohort study in cardiovascular diseases, and machine learning for pretest probability of obstructive coronary stenosis in symptomatic patients. He was a recipient of the 2012 China National Scholarship, the 2016 Science and Technology Progress Award of the Ministry of Education, the 2017 SCCT Young Researchers Award, and the 2017 Chinese Medical Association Radiology Spark Program Award. He has been published a total of 48 papers, including 16 SCI papers.

CHEN CHEN received the bachelor's degree in the Internet of Things engineering from the Harbin Institute of Technology, China, the master's degree in advanced computing from Imperial College London, in 2016, and an Honor Graduate, in 2015. She is currently pursuing the Ph.D. degree with Imperial College London, under the supervision of Prof. D. Rueckert.

She was a Deep Learning Algorithm Engineer and a Team Leader with Infervision, Beijing, for one year. She had worked on several projects that aimed at developing deep learning algorithms to analyze medical CT images automatically. She is also currently involved in the SmartHeart Programme, U.K. Her current research interests include medical image analysis, deep learning, and computer vision.

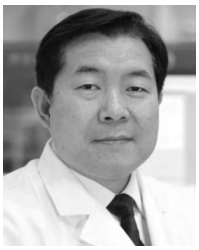
During the bachelor's degree, she was a recipient of the China National Scholarship, in 2013 and 2014.

ZHI HAO born in Xinzhou, Shanxi, China, in 1993. He received the M.S. degree in communication and information system from Yanshan University, Hebei, China, in 2018.

From 2017 to 2018, he was a Deep Learning Algorithm Intern with Infervision, Beijing, for one year. He had worked on several projects that aimed at developing deep learning algorithms to analyze medical CT images automatically.

YUNQIANG AN was born in Jiamusi, Heilongjiang, China, in 1986. He received the B.S. degree in theoretical and applied mechanics from Jilin University, Changchun, China, in 2011, and the Ph.D. degree in biomedical engineering from Capital Medical University, Beijing, China, in 2017. He is currently an Engineer of the Department of Radiologic Imaging, Fuwai Hospital, China. His research interest includes the high simulation biomechanical model and application, focusing on respiratory systems and cardiovascular systems. He was a recipient of the First Prize of the Eighth National Zhou Peiyuan Mechanical Competition, Jilin Province Division, and the China National Scholarship.

SEN LIANG was born in Taihe, Jiangxi, China, in 1992. He received the B.S. degree from the College of Computer Science and Technology, Jilin University, Changchun, China, in 2014, where he is currently pursuing the Ph.D. degree under the supervision of Prof. Y. Xu. His current research interests include medical image analysis, deep learning, and computer vision.



BIN LU received the B.S.Med. degree in clinical medicine from Taishan Medical University, Taishan, China, and the M.D. degree in clinical medicine from Chinese Academy of Medical Science, Peking Union Medical College, Beijing, China, in 1998. He held a postdoctoral position with the Harbor-UCLA Medical Center, Torrance, CA, USA.

He is currently a Professor and the Director of the Department of Radiologic Imaging, Fuwai Hospital, China. He has authored four books and more than 160 articles. His research interests include multidisciplinary research approach to systematically assess and validate the clinical utility of novel cardiovascular computer tomography imaging technologies, including coronary artery disease, atherosclerosis, and cardiac function.

Dr. Lu was a recipient of the Commission of Cardiovascular and Thoracic Committee of Chinese Society of Radiology, the Chairman of Society of Cardiovascular Computed Tomography China IRC, the Vice-President of the Asian Society of Cardiovascular Imaging, and the Gold Medal Award of Asian Society of Cardio-vascular Imaging.