Check for updates

OPEN ACCESS

EDITED BY Lifang Zhou, Chongqing University of Posts and Telecommunications, China

REVIEWED BY Artur Yakimovich, Helmholtz Association of German Research Centers (HZ), Germany Imran Iqbal, New York University, United States

*CORRESPONDENCE Yingwei Guo ☑ guoyingwei8801@163.com Huai Chen ☑ chenhuai1977@163.com Yan Kang ☑ kangyan@sztu.edu.cn

[†]These authors have contributed equally to this work

RECEIVED 18 July 2024 ACCEPTED 21 February 2025 PUBLISHED 12 March 2025

CITATION

Yang Y, Zheng J, Guo P, Gao Q, Guo Y, Chen Z, Liu C, Wu T, Ouyang Z, Chen H and Kang Y (2025) Three-stage registration pipeline for dynamic lung field of chest X-ray images based on convolutional neural networks. *Front. Artif. Intell.* 8:1466643. doi: 10.3389/frai.2025.1466643

COPYRIGHT

© 2025 Yang, Zheng, Guo, Gao, Guo, Chen, Liu, Wu, Ouyang, Chen and Kang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Three-stage registration pipeline for dynamic lung field of chest X-ray images based on convolutional neural networks

Yingjian Yang^{1†}, Jie Zheng^{1†}, Peng Guo^{1†}, Qi Gao², Yingwei Guo^{3*}, Ziran Chen⁴, Chengcheng Liu⁵, Tianqi Wu¹, Zhanglei Ouyang¹, Huai Chen^{6*} and Yan Kang^{4,7,8*}

¹Department of Radiological Research and Development, Shenzhen Lanmage Medical Technology Co., Ltd., Shenzhen, Guangdong, China, ²Neusoft Medical System Co., Ltd., Shenyang, Liaoning, China, ³School of Electrical and Information Engineering, Northeast Petroleum University, Daqing, China, ⁴College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China, ⁵School of Life and Health Management, Shenyang City University, Shenyang, China, ⁶Department of Radiology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, ⁷College of Health Science and Environmental Engineering, Shenzhen Technology University, Shenzhen, China, ⁸Engineering Research Centre of Medical Imaging and Intelligent Analysis, Ministry of Education, Shenyang, China

Background: The anatomically constrained registration network (AC-RegNet), which yields anatomically plausible results, has emerged as the state-of-the-art registration architecture for chest X-ray (CXR) images. Nevertheless, accurate lung field registration results may be more favored and exciting than the registration results of the entire CXR images and hold promise for dynamic lung field analysis in clinical practice.

Objective: Based on the above, a registration model of the dynamic lung field of CXR images based on AC-RegNet and static CXR images is urgently developed to register these dynamic lung fields for clinical quantitative analysis.

Methods: This paper proposes a fully automatic three-stage registration pipeline for the dynamic lung field of CXR images. First, the dynamic lung field mask images are generated from a pre-trained standard lung field segmentation model with the dynamic CXR images. Then, a lung field abstraction model is designed to generate the dynamic lung field images based on the dynamic lung field mask images and their corresponding CXR images. Finally, we propose a three-step registration training method to train the AC-RegNet, obtaining the registration network of the dynamic lung field images (AC-RegNet_V3).

Results: The proposed AC-RegNet_V3 with the four basic segmentation networks achieve the mean dice similarity coefficient (DSC) of 0.991, 0.993, 0.993, and 0.993, mean Hausdorff distance (HD) of 12.512, 12.813, 12.449, and 13.661, mean average symmetric surface distance (ASSD) of 0.654, 0.550, 0.572, and 0.564, and mean squared distance (MSD) of 559.098, 577.797, 548.189, and 559.652, respectively. Besides, compared to the dynamic CXR images, the mean DSC of these four basic segmentation networks with AC-RegNet has been significantly improved by 7.2, 7.4, 7.4, and 7.4% (*p*-value < 0.0001). Meanwhile, the mean HD has been significantly improved by 8.994, 8.693, 9.057, and 7.845 (*p*-value < 0.0001). Similarly, the mean ASSD has significantly improved by 4.576, 4.680, 4.658, and 4.658 (*p*-value < 0.0001). Last, the mean MSD has significantly improved by 508.936, 519.776, 517.904, and 520.626 (*p*-value < 0.0001).

Conclusion: Our proposed three-stage registration pipeline has demonstrated its effectiveness in dynamic lung field registration. Therefore, it could become a powerful tool for dynamic lung field analysis in clinical practice, such as pulmonary airflow detection and air trapping location.

KEYWORDS

dynamic chest X-ray images, lung field segmentation, medical image registration, anatomical constraints, convolutional neural network, AC-RegNet

1 Introduction

Compared with computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), PET-CT, and other imaging devices, X-ray is the most widely used primary chest imaging technique as it is widely available, low-cost, fast imaging speed, and easy to acquire (Howell, 2016; Seah et al., 2021; Yang et al., 2024). Notably, its characteristic of fast imaging speed (seconds after exposure) makes the X-ray the preferred chest imaging device to improve work efficiency and facilitate the diagnosis of routine, critically ill, and emergency settings in clinical chest imaging examinations (Yang et al., 2024; Yang et al., 2024).

Although traditional static chest X-ray (CXR) images can display gross lesions such as lung inflammation, lung lumps, tuberculosis, etc., they lack corresponding information on dynamic lung motion. Therefore, dynamic CXR images during free breathing are captured to analyze dynamic indicators during lung respiration, such as hemidiaphragm motion (Yang et al., 2024; Chen et al., 2022) and dynamic cardiothoracic ratio detection (Yang et al., 2024; Jafar et al., 2022). However, for dynamic CXR images, detecting the above dynamic indicators is not enough in clinical practice. It is more important to analyze further the detected dynamic indicators to provide reasonable clinical recommendations. Like the parametric response mapping in expiratory and inspiratory chest CT images (Deng et al., 2024; Wang et al., 2024), registration technology is indispensable to accurately determine the point-to-point relationship between the above dynamic indicators of these dynamic CXR images.

Based on the above, medical image registration technology is a crucial step and pillar problem in medical image analysis for aligning the source image (moving image) with the target image (fixed image) (Mansilla et al., 2020). Essentially, registration technology aims to find the deformation fields of the point-to-point correspondence between the source and target image. Specifically, the classic unsupervised technology for medical image registration, SimpleElastix (Marstal et al., 2016), is widely used in non-deformable tissues such as threedimensional (3D) brain and two-dimensional (2D)/3D bone X-ray/ CT/MRI images (Guo et al., 2022; Guo et al., 2022; Guo et al., 2022; Yang and Guo, 2024; Chang and Lee, 2021). Compared with SimpleElastix, the end-to-end deep convolutional neural network is constantly proposed for the intelligent processing of medical images (Iqbal et al., 2021; Iqbal et al., 2025; Yang et al., 2022). However, it is disappointing that most of the supervised registration techniques for chest medical images are focused on 3D CT images (Xiao et al., 2023), and only a few studies have been conducted on 2D X-ray images from 2016 to the present (Nie et al., 2024). The only registration method, AC-RegNet, for chest CXR images during this period was proposed in 2020 (Mansilla et al., 2020). Specifically, this technique produces anatomically plausible results for 2D chest CXR images. The evaluation metrics of the proposed AC-RegNet surpass the AE-RegNet, CE-RegNet, RegNet, and SimpleElastix (Mansilla et al., 2020; Marstal et al., 2016). Therefore, the AC-RegNet has been considered the stateof-the-art registration architecture for chest X-ray (CXR) images. Nevertheless, the above models have not achieved registration of dynamic lung field images collected during free or forced breathing. Therefore, this limits the quantitative analysis of the dynamic lung field based on the CXR images.

Accurate lung field registration results may be more favored and exciting than the registration results of the entire CXR images for dynamic lung field analysis in clinical practice, such as pulmonary airflow detection (Jiang et al., 2024) and air trapping location (Zhang et al., 2024). Specifically, the lung undergoes a non-rigid and complex process of contraction and expansion during breathing (Santhanam, 2006). Compared to other static organs or tissues (such as the brain, bones, etc.), these non-rigid and complex deformations present significant challenges for lung field registration. Registration is even more challenging for dynamic chest CXR images projected in 2D further. In addition, the limited quantity and quality of dynamic CXR images also constrain the development of dynamic lung field registration. Despite the existing problems and challenges, we still hope to propose a registration model for dynamic lung field images based on the limited quantity and quality of static CXR images to expand the clinical application of dynamic chest CXR images and improve the accuracy of quantitative analysis based on lung fields.

However, the limited number of dynamic CXR images makes it challenging to train the AC-RegNet, which limits the clinical quantitative analysis of dynamic lung fields. Based on the above, we propose a fully automatic registration pipeline for the dynamic lung field of CXR images based on AC-RegNet architecture and static posteroanterior (P-A) CXR images. First, the dynamic lung field mask images are generated from a pre-trained standard lung field segmentation model with the dynamic P-A CXR images. Then, a lung field abstraction model is designed to generate the dynamic lung field images based on the dynamic lung field mask images and their corresponding P-A CXR images. Finally, we propose a three-step registration training method to train the basic architecture (VectorCNN and differentiable warper) of AC-RegNet, obtaining the registration network of the dynamic lung field images. Our contributions in this paper are briefly described as follows:

- (1) Overall, this paper presents a fully automatic three-stage registration pipeline for the dynamic lung field of CXR images based on AC-RegNet, which effectively addresses the issue of the inability to register dynamic lung field images for the quantitative analysis of lung fields.
- (2) Essentially, this paper proposes a three-step registration training method that includes initial training, enhanced

training, and final training of the basic architecture of AC-RegNet based on static P-A CXR images, which fully utilizes static P-A CXR images to address the registration issue of these non-rigid and complex deformations in the dynamic lung field.

(3) Clinically, the proposed fully automatic registration pipeline can effectively register dynamic lung field images to maintain lung field morphology alignment during respiration. Thus, this may become a valuable tool for further quantitative analysis of dynamic lung fields, such as pulmonary airflow detection, air trapping location, etc.

2 Materials and methods

2.1 Materials

Figure 1 intuitively shows the detailed distribution of the static and dynamic P-A CXR images in the six datasets. Meanwhile, Figure 1a intuitively shows each dataset's proportion of normal, abnormal, and unclear cases. Besides, Figure 1b intuitively shows the proportion of pneumonia and tuberculosis cases in each dataset.

Specifically, five datasets, D1-D5, include 786 sets of static 512×512 P-A CXR images, and dataset D6 includes 11 sets of dynamic 512 × 512 CXR images. These five datasets, D1-D5, are from publicly available datasets (Yang et al., 2024). Besides, 11 sets of dynamic P-A CXR images of dataset D6 are selected from the case of CXR video. This case of CXR video collected from a female participant aged 53 during free breathing using a digital X-ray imaging system (manufacturer: Lanmage, collection mode: sequence point slice, exposure parameters: 78KV, 200 mA, 50 ms, and flat panel detector: IRAY) for chest photography. Besides, due to the unstable radiation dose during the startup of the digital X-ray imaging system, the 13 video frames and subsequent frames in the video image tend to stabilize. Therefore, these 11 sets of dynamic 3,072 × 3,072 P-A CXR images are 13-23 video frames extracted from the CXR video. Then, to ensure that these dynamic P-A CXR images are the same size as static P-A CXR images, resize them to $512 \times 512.$

This female participant was provided written informed consent, and the Guangzhou Medical University Ethics Committee in China approved this study (Grant number: 2023-hg-ks-24, Approval Date: 28 August 2023, Tel: +86-20-34153599, and Fax: +86-20-34153066).

2.2 Methods

Figure 2 shows the three-stage structure of the dynamic lung field registration pipeline of CXR images based on AC-RegNet architecture. First, the dynamic lung field mask images of the CXR images are generated by a pre-trained standard and pathological lung field segmentation network. Then, the dynamic lung field mages are abstracted based on the dynamic lung field mask images and their corresponding CXR images. Lastly, dynamic lung field images can be registered at different times based on the pre-trained AC-RegNet (registration network). The source code is available on the website: https://github.com/YingjianYang/Dynamic-Lung-Field-Registration.



FIGURE 1

Data distribution map of P-A CXR images in each dataset; (a) Case classification map of P-A CXR images in each dataset; (b) Abnormal case classification map of P-A CXR images in each dataset.



2.2.1 Stage I: dynamic lung field segmentation

A cross-center and standard segmentation network of pathological lungs is crucial for extracting lung fields from dynamic P-A CXR images, effectively assisting the registration of dynamic lung field images.

Convolutional neural networks (CNNs) have become the leading technical means for medical image segmentation tasks (Yang et al., 2022; Long et al., 2015; Badrinarayanan et al., 2017; Ronneberger et al., 2015; Yang et al., 2021; Jha et al., 2019; Wang et al., 2021; Zaman et al., 2024; Zeng et al., 2023; Duan et al., 2023), providing excellent technical support for registering dynamic lung fields. Our previous research separately utilized these cross-center 755 static P-A CXR mages and their lung field label images to train five basic CNNs, achieving corresponding cross-center and standard pre-train segmentation networks (Yang et al., 2024). These five CNNs include FCN (Long et al., 2015), SegNet (Badrinarayanan et al., 2017), U-Net (Ronneberger et al., 2015), ResU-Net++ (Yang et al., 2021; Jha et al., 2019), and AttU-Net (Wang et al., 2021). However, the above research found that due to the lack of skip connections between each level in FCN, the pre-train lung field segmentation network based on FCN lacks detailed information and exhibits noticeable jagged edges at the edges of the lung field mask. Therefore, the other four pre-trained lung field segmentation networks can be used to perform lung field segmentation tasks for dynamic P-A CXR images.

2.2.2 Stage II: dynamic lung field extraction

Dynamic lung field images are extracted from the dynamic P-A CXR images to register dynamic lung field images. Like the lung field extraction of the chest CT images for calculating lung parenchyma parameters (Yang et al., 2021), locating chronic obstructive pulmonary disease (Yang et al., 2020), and calculating lung radiomics feature (Yang et al., 2022), the lung field extraction of dynamic P-A CXR images is based on the dynamic lung field mask images generated from the segmentation network and their CXR images.

The pixel values in the lung field of the dynamic lung field mask images should be set to 1 instead of the pixel values in the non-lung field to 0, obtaining the dynamic preprocessed lung field mask images. Then, the dynamic preprocessed lung field mask images are multiplied by their CXR images at the pixel level to obtain the dynamic lung field images. The above pixel values set maintain the grayscale of the dynamic lung field mask images consistent with the grayscale of the lung field in the CXR image.

2.2.3 Stage III: dynamic lung field registration

Since the AC-RegNet architecture was proposed, this technique has produced anatomically plausible results and has been considered the state-of-the-art registration architecture for CXR images (Mansilla et al., 2020; Nie et al., 2024). In this work, we extend the application of the AC-RegNet architecture to the registration of dynamic lung field images. Specifically, this basic architecture of AC-RegNet includes two main modules: VectorCNN and differentiable warper (Mansilla et al., 2020; Jaderberg et al., 2015).

Figure 3 shows the training process of the registration network (AC-RegNet), including the encoder and basic architecture training. Specifically, we follow the previous training method of AC-RegNet. First, 787 static CXR images and their lung field label images were used to train the anatomically constrained network [a kind of denoising autoencoders (Vincent et al., 2010)] on the lung field. Then, the encoder module of this pre-train anatomically constrained network participates in the training of VectorCNN and differentiable warper in the basic architecture of AC-RegNet. Since the registration task is to achieve dynamic lung field registration, the heart labels are discarded during the training of this anatomically constrained network.

Specifically, we propose a three-step registration training method to train the VectorCNN and differentiable warper of AC-RegNet. The three cascaded steps separately complete the initial, enhanced, and final training of the basic architecture of AC-RegNet based on 787 static P-A CXR images. First, any two pairs of these 787 static P-A CXR images (787 × 787–787 = 618,582) are combined. Then, they are configured as the source and target images for initially training the VectorCNN and differentiable warper (Step 2.1: Initial training process of VectorCNN and differentiable warper). Second, 787 static affine P-A CXR images are generated by affine transformation [a kind of data augmentation technique (Yang et al., 2024; Chlap et al., 2021)] of static P-A CXR images, respectively. Subsequently, 787 pairs of the source and target images are configured based on these 787 static P-A CXR images, and their affine images are used for enhanced training of the initial pre-trained VectorCNN and differentiable warper



(Step 2.2: Enhanced training process of VectorCNN and differentiable warper). Third, 787 pairs of static lung field images are abstracted from their static P-A CXR images using the dynamic lung field extraction method in Section 2.2.2. Then, 787 static affine lung field images are generated by the affine transformation of 787 pairs of static lung field images. Similarly, these static lung field images and their affine images are used for the final training of the enhanced pre-trained VectorCNN and differentiable warper (Step 2.3: Final training process of VectorCNN and differentiable warper). Notably, AC-RegNet introduces an anatomically constrained network to produce anatomically plausible results during training. Specifically, the 618,582 $(787 \times 787 - 787)$ pairs of static lung field mask images participate in the initial training (Step 2.1 AC-RegNet_V1). Similarly, 787 pairs of lung field mask images and their affine images participate in the enhanced and final training (Step 2.2 AC-RegNet_V2 and Step 2.3 AC-RegNet_V3), respectively. It should be noted that the randomized parameters of affine transformation to each affine P-A CXR/lung field image and its corresponding affine lung field mask image should be consistent.

2.2.4 Loss function

During each training step, the loss values are meticulously calculated to adjust the VectorCNN and differentiable warper parameters based on the following comprehensive loss function $Loss_t(I_t,J_t,S_{It},S_{Jt},T_t)_{loss}$ shown in Equation 1 (Yang et al., 2024; Mansilla et al., 2020; Balakrishnan et al., 2018; Ferrante et al., 2018; Yeap et al., 2023).

$$Loss_{t}(I_{t}, J_{t}, S_{lt}, S_{Jt}, T_{t})_{loss}$$

$$= -\lambda_{NCC} \bullet NCC_{t}(I_{t} \circ T_{t}, J_{t})_{loss} + \lambda_{TV} \bullet TV_{t}(T_{t})_{loss}$$

$$+\lambda_{BCE} \bullet BCE_{t}(S_{lt} \circ T_{t}, S_{Jt})_{loss}$$

$$+\lambda_{L2} \bullet L2_{t}(enc(S_{lt} \circ T_{t}), enc(S_{Jt}))_{loss}$$
(1)

Where train step t = 1,2,3 represents the initial, enhanced, and final training process of VectorCNN and differentiable warper, respectively; parameters I_t , J_t , S_{lb} , S_{lb} , and T_t separately represent the source image, target image, source mask, target mask, and deformation field at the *t*th train step; functions NCC_t (), TV_t (), BCE_t (), and $L2_t$ () separately represent the Negative normalized cross-correlation (NCC) loss (Balakrishnan et al., 2018), total variation (TV) loss (Ferrante et al., 2018), binary cross-entropy (BCE) Loss (Yang et al., 2024; Mansilla et al., 2020), and L2 loss (Mansilla et al., 2020) at the *t*th train step. Meanwhile, the weight factors of functions NCC_t (), TV_t (), BCE_t (), and $L2_t$ () are set to the previous default values. Therefore, the weight factors are: $\lambda_{NCC}=1.0$, $\lambda_{TV}=5.0 \times 10^{-5}$, $\lambda_{BCE}=1.0$, λ_{L2} = 1.0×10^{-1} . Last, *enc* represents the pre-trained encoder.

Specifically, Equation 2 presents the precise mathematical expression of NCC loss, which accurately describes the degree of correlation between the source images and their target images (Balakrishnan et al., 2018).

$$-NCC_t \left(I_t \circ T_t, J_t \right)_{loss} = -\frac{\operatorname{cov} \left(I_t \circ T_t, J_t \right)}{\sqrt{Var(I_t \circ T_t) \cdot Var(J_t)} + \varepsilon}$$
(2)

Where a non-zero factor ε is added to ensure that the denominator of NCC loss is not zero ($\varepsilon = 1 \times 10^{-5}$), and functions *Cov* () and *Var* () separately represent the covariance and variance calculation.

$$\operatorname{cov}(I_t \circ T_t, J_t) = E\Big[\big(I_t \circ T_t - E\big(I_t \circ T_t\big)\big)\big(J_t - E\big(J_t\big)\big)\Big]$$
(3)

$$Var(I_t \circ T_t) = E\left[\left(I_t \circ T_t - E(I_t \circ T_t)\right)^2\right]$$
(4)

$$Var(J_t) = E\left[\left(J_t - E(J_t)\right)^2\right]$$
(5)

Where function E () represents the expected value calculation. Substitute these Equations 3–5 into Equation 2, obtaining Equation 6.

$$-NCC_{t}(I_{t} \circ T_{t},J_{t})_{loss} = -\frac{E\begin{bmatrix} (I_{t} \circ T_{t} - E(I_{t} \circ T_{t})) \\ (J_{t} - E(J_{t})) \end{bmatrix}}{\sqrt{E\begin{bmatrix} (I_{t} \circ T_{t} - E(I_{t} \circ T_{t}))^{2} \end{bmatrix}} \cdot \\ +\varepsilon}$$

$$E\begin{bmatrix} (J_{t} - E(J_{t}))^{2} \end{bmatrix}$$
(6)

The final simplified Equation 7 can be obtained using the expected linear property.

$$-NCC_{t}(I_{t} \circ T_{t}, J_{t})_{loss} = -\frac{E\begin{bmatrix} (I_{t} \circ T_{t} - E(I_{t} \circ T_{t})) \\ (J_{t} - E(J_{t})) \end{bmatrix}}{\sqrt{\left[E((I_{t} \circ T_{t}^{2}) - (E(I_{t} \circ T_{t}))^{2}\right]^{\bullet}} + \varepsilon} + \varepsilon \quad (7)$$

The simplification process from Equations 6, 7 is shown in Equations 8–10.

$$-NCC_{t}(I_{t} \circ T_{t}, J_{t})_{loss} = -\frac{E[(I_{t} \circ T_{t} - E(I_{t} \circ T_{t}) - E(I_{t} \circ T_{t})]}{\left[\frac{E[((I_{t} \circ T_{t})^{2} - 2(I_{t} \circ T_{t}) \bullet - E(I_{t} \circ T_{t})]^{2}] \bullet}{\sqrt{\frac{E[(J_{t} \circ T_{t}) + (E(I_{t} \circ T_{t}))^{2}] \bullet}{\left[\frac{E[(J_{t}^{2} - 2J_{t} \bullet E(J_{t}) + E(J_{t} \circ T_{t})]^{2}] \bullet} + \varepsilon}}\right]}}$$
(8)

$$-NCC_{t}(I_{t} \circ T_{t}, J_{t})_{loss} = -\frac{E[(I_{t} \circ T_{t} - E(I_{t} \circ T_{t}))]}{\left[E((I_{t} \circ T_{t})^{2}) - 2E((I_{t} \circ T_{t}) \bullet \\ E(I_{t} \circ T_{t})) + E(E(I_{t} \circ T_{t}))^{2}\right] \bullet} + \varepsilon$$

$$\left[E(J_{t}^{2}) - 2E(J_{t} \bullet E(J_{t})) + \\ E(E(J_{t}))^{2}\right]$$
(9)

$$-NCC_{t}(I_{t} \circ T_{t}, J_{t})_{loss} = -\frac{E[(I_{t} \circ T_{t} - E(I_{t} \circ T_{t}))]}{\left[E((I_{t} \circ T_{t})^{2}) - 2(E(I_{t} \circ T_{t}))^{2} + (E(I_{t} \circ T_{t}))^{2}] \bullet [E(J_{t}^{2}) - 2 + \varepsilon (E(J_{t}))^{2}]\right]}$$
(10)

Specifically, Equation 11 shows the mathematical expression of TV loss, which smooths the source images' registration images (Ferrante et al., 2018).

$$TV(T_{i})_{loss} = \begin{bmatrix} mean(TV(T_{ix})_{loss}) + \\ mean(TV(T_{iy})_{loss}) \end{bmatrix} / 2$$
$$= \begin{bmatrix} mean\left(\sum_{i,j}^{n_{i}} |t_{i+1,j} - t_{i,j}|\right) + \\ mean\left(\sum_{i,j}^{n_{2}} |t_{i,j+1} - t_{i,j}|\right) \end{bmatrix} / 2$$
(11)
$$= \begin{bmatrix} \left(\sum_{i,j}^{n_{1}} |t_{i+1,j} - t_{i,j}|\right) / n_{1} + \\ \left(\sum_{i,j}^{n_{2}} |t_{i,j+1} - t_{i,j}|\right) / n_{2} \end{bmatrix} / 2$$

Where function *mean* () represents the average operation and this function $|\bullet|$ represents the absolute value operation; deformation field $\begin{pmatrix} t_{11} & \dots & t_{1n_i} \end{pmatrix}$

 $T_t = \begin{bmatrix} \vdots & \vdots & \vdots \\ t_{n_21} & \cdots & t_{n_n} \\ t_{n_21} & \cdots & t_{n_n} \end{bmatrix}, t_{i,j} \subset T_t; \text{ n1 and n2 are the dimensions of}$

deformation field T_t at the *t*th train step; T_{tx} and T_{ty} separately represent the element of the x-axis and y-axis of the dimension of deformation field T_t at the *t*th train step.

Specifically, Equation 12 shows the mathematical expression of BCE loss for calculating the difference between these target mask images and the warped mask images generated by the source mask images and differentiable warper (Yang et al., 2024; Mansilla et al., 2020).

$$BCE_t \left(S_{It} \circ T_t, S_{Jt} \right)_{loss} = \frac{1}{N} \sum_{i} - \begin{bmatrix} S_{Jti} \cdot \log(S_{Iti} \circ T_{ti}) + \\ (1 - S_{Jti}) \cdot \\ \log(1 - S_{Iti} \circ T_{ti}) \end{bmatrix}$$
(12)

Where S_{Jti} , S_{Jti} , T_{ti} , and N separately represent the *i*th source mask, target mask, deformation field at the *t*th train step, and the number of *i*.

Specifically, Equation 13 shows the mathematical expression of L2 loss for calculating the difference between these encoding target mask images and their encoding warped mask images (Mansilla et al., 2020).

$$L2_{t}\left(enc\left(S_{It}\circ T_{t}\right),enc\left(S_{Jt}\right)\right)_{loss} = mean\left(\left\|\frac{enc\left(S_{It}\circ T_{t}\right)-}{enc\left(S_{Jt}\right)}\right\|_{2}^{2}\right)$$

$$=\sum_{j}^{n}\left(\left(\sum_{i,j=1}^{n}\left(st_{i,j}-s_{i,j}\right)^{2}\right)/n\right)$$
(13)

Where function *mean* () represents the average operation and this function $\| \cdot \|_2$ represents the 2-norm operation; the encoding warped $\begin{pmatrix} st_{11} & \dots & st_{1n} \end{pmatrix}$

mask images $enc(S_{It} \circ T_t) = \begin{pmatrix} st_{11} & \dots & st_{1n} \\ \vdots & \ddots & \vdots \\ st_{n1} & \dots & st_{nn} \end{pmatrix}$, $st_{i,j} \subset enc(S_{It} \circ T_t)$; the encoding target mask images $enc(S_{Jt}) = \begin{pmatrix} s_{11} & \dots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \dots & s_{nn} \end{pmatrix}$,

 $s_{i,j} \subset enc(S_{Jt})$; $n \times n$ is the size of the encoding warped mask images or target mask images.

Finally, this final-train network is applied to register the dynamic lung field images of 512×512 P-A CXR images in the dataset D6.

3 Experiments

This section conducts comprehensive ablation and comparison studies based on the above materials and methods. The purpose of these studies is to reflect the scientific validity and superior performance of our proposed registration model (AC-RegNet_V3) in a variety of scenarios.

3.1 Experimental design

Figure 4 shows the experimental design for the ablation study to compare the effects of different steps (AC-RegNet_V1, AC-RegNet_V2, AC-RegNet_V3). Specifically, the fixed lung filed mask images are generated from the dynamic P-A CXR images based on a pre-trained lung field segmentation network [U-Net (Ronneberger et al., 2015)]. Besides, experiments 1 and 2 register any two pairs of these 11 dynamic P-A CXR images. Unlike experiments 1 and 2, experiment 3 registers any two pairs of these dynamic lung field images abstracted from the dynamic P-A CXR images based on their lung field mask images.

In addition, Figure 5 shows the experimental design for the comparing study to compare the performance of our proposed registration model and the SimpleElastix. Previous research confirmed that the evaluation metrics of the AC-RegNet have surpassed the AE-RegNet, CE-RegNet, and RegNet (Mansilla et al., 2020). Therefore, the evaluation metrics among these registration models above will not be compared again in the experimental section. Specifically, four pre-trained lung field segmentation networks, SegNet (Badrinarayanan et al., 2017), U-Net (Ronneberger et al., 2015)], ResU-Net++ (Yang et al., 2021; Jha et al., 2019), and AttU-Net (Wang et al., 2021), to compare the effects of different lung field segmentation networks on the proposedAC-RegNet_V3. Besides, the widely used unsupervised technology for medical image registration, SimpleElastix (Marstal et al., 2016) with the affine transform, is also compared with the proposed AC-RegNet_V3. Specifically, these four pre-trained lung field segmentation networks are separately applied to segment the lung fields of 11 sets of dynamic P-A CXR images, obtaining four sets of dynamic lung field mask images in Stage I. Then, four groups of dynamic lung field images are abstracted from the dynamic P-A CXR images based on their lung field mask images in stage II. Ultimately, 110 pairs of the source and target images are generated by combining any two pairs of these 11 sets of dynamic lung field images in each group. Last, the AC-RegNet_V3 and SimpleElastix are applied separately to register the dynamic lung field image groups in stage III.





3.2 Evaluation metrics

Our previous study has displayed the evaluation metrics of these four pre-trained segmentation networks (Yang et al., 2024; Yang et al., 2024). Therefore, this section mainly presents and compares the registration evaluation metrics of these dynamic CXRs, lung fields, and registration images.

After completing the registration of 110 pairs of the source and target images, four stand evaluation metrics, dice similarity coefficient (DSC) (Mansilla et al., 2020; Yeap et al., 2023), Hausdorff distance

(HD) (Yang et al., 2024), average symmetric surface distance (ASSD) (Mansilla et al., 2020) and mean squared distance (MSD) (Rahunathan et al., 2005), are used to evaluate these dynamic CXR, lung field, and registration images. Besides, unlike medical images of non-deformable tissues such as brain and bone images (Guo et al., 2022; Guo et al., 2022; Guo et al., 2022; Yang and Guo, 2024; Chang and Lee, 2021; Guo et al., 2022), the lung fields in dynamic CXR images undergo anatomical folding during respiration. Therefore, the Jacobian determinant (Kuang, 2019) is excluded from this study to evaluate the registration performance.

The DSC, HD, and ASSD focus on changes in lung field morphology based on their mask, while the evaluation metric MSD focuses on changes in the CXR or lung field image. Equations 14–17 show the mathematical expression of these four stand evaluation metrics (Yang et al., 2024; Mansilla et al., 2020; Yeap et al., 2023; Rahunathan et al., 2005).

$$DSC = 2 \times \frac{\left|I_{fixed} \cap I_{registered/moving}\right|}{\left|I_{fixed}\right| + \left|I_{registered/moving}\right|}$$
(14)

$$HD = \max\left(d_{I_{fixed}}I_{registered/moving}, d_{I_{registered/moving}}I_{fixed}\right)$$
$$= \max\left\{\begin{cases}\max_{x \in I_{fixed}} y \in I_{registered/moving}\\\max_{y \in I_{registered/moving}} x \in I_{fixed} \end{cases}\right\}$$
(15)

$$ASSD = \frac{1}{\left|I_{fixed}\right| + I_{registered/moving}} \times \left(\sum_{\substack{(x,y) \in I_{fixed} \\ \sum \\ (x,y) \in I_{registered/moving}}} \|(x,y), I_{fixed}\| + \sum_{\substack{(x,y) \in I_{registered/moving}}} \|(x,y), I_{registered/moving}\| \right)$$
(16)

$$MSD = \frac{1}{N} \sum_{(x,y)} \left[I_{fixed} \left(x, y \right) - I_{registered/moving} \left(x, y \right) \right]^2$$
(17)

TABLE 1 The evaluation metrics of the ablation study.

Where I_{fixed} represents the fixed lung filed mask image of the $I_{fixed}(x,y)$, and $I_{fixed}(x,y)$ represents the fixed CXR image. Besides, $I_{registered/moving}$ represents the registered/moving lung filed mask image of $I_{registered/moving}(x,y)$, and $I_{registered/moving}(x,y)$ represents

the registered/moving CXR image. Last, this function $\|$ represents the 2-norm operation, and *N* represents the pixel number of the fixed or registered/moving CXR image.

4 Results

This section comprehensively compares and presents the registration results of dynamic lung field images based on the above materials and methods.

4.1 Ablation study results

Table 1 separately shows the descriptive statistics on the evaluation metrics of the 110 pairs of images based on the ablation study. Besides, Figure 6 shows the visual distribution of the evaluation metrics of the Baseline and AC-RegNet_V1-V3 in the ablation study, respectively.

Specifically, DSC (mean \pm SD), HD (mean \pm SD), ASSD (mean \pm SD), and MSD (mean \pm SD) of these 110 pairs of dynamic CXR images (Baseline in Figure 6) are 0.919 \pm 0.047, 21.506 \pm 8.243, 5.230 \pm 2.947, and 612.654 \pm 503.383, respectively. Meanwhile, the evaluation metrics of the 110 pairs of dynamic CXR images based on AC-RegNet_V1 achieve the DSC (mean \pm SD) of 0.909 \pm 0.049, HD

Evaluation metrics	Network			110 pairs of images
	Segmentation	Lung field extraction	Registration	
DSC (mean ± SD*)	×	×	×	$0.919~(0.808{-}0.987)\pm0.047$
	×	×	AC-RegNet_V1	$0.909~(0.762{-}0.980)\pm0.049$
	×	×	AC-RegNet_V2	0.933 (0.840-0.989) ± 0.034
	U-Net [26]	1	AC-RegNet_V3	0.993 (0.901-1.000) ± 0.016
HD (mean ± SD*)	×	×	×	21.506 (5.385-39.000) ± 8.243
	×	×	AC-RegNet_V1	26.074 (6.083-43.046) ± 7.577
	×	×	AC-RegNet_V2	20.755 (4.472-39.000) ± 7.642
	U-Net [26]	1	AC-RegNet_V3	12.813 (1.000-42.202) ± 11.460
ASSD (mean ± SD*)	×	×	×	5.230 (0.866-12.130) ± 2.947
	×	×	AC-RegNet_V1	5.197 (1.245-10.690) ± 2.253
	×	×	AC-RegNet_V2	4.243 (0.719-9.270) ± 2.116
	U-Net [26]	1	AC-RegNet_V3	0.550 (0.035-4.361) ± 0.911
MSD (mean ± SD*)	x	×	×	612.654 (32.924– 2050.390) ± 503.383
	×	×	AC-RegNet_V1	250.506 (46.376– 926.845) ± 212.108
	×	×	AC-RegNet_V2	422.766 (23.824– 1688.810) ± 381.697
	U-Net [26]	1	AC-RegNet_V3	92.878 (9.438-616.990) ± 106.752

* SD, standard deviation.



(mean ± SD) of 26.074 ± 7.577 , and ASSD (mean ± SD) of 5.197 ± 2.253 , and MSD (mean ± SD) of 250.506 ± 212.108 . The evaluation metrics of the 110 pairs of dynamic CXR images based on AC-RegNet_V2 achieve the DSC (mean ± SD) of 0.933 ± 0.034 , HD (mean ± SD) of 20.755 ± 7.642 , and ASSD (mean ± SD) of 4.243 ± 2.116 , and MSD (mean ± SD) of 422.766 ± 381.697 . The evaluation metrics of the 110 pairs of dynamic CXR images based on AC-RegNet_V3 achieve the DSC (mean ± SD) of 0.993 ± 0.016 , HD (mean ± SD) of 12.813 ± 11.460 , and ASSD (mean ± SD) of 0.550 ± 0.911 , and MSD (mean ± SD) of 92.878 ± 106.752 .

Compared with the evaluation metrics of the Baseline, only all the evaluation metrics of the AC-RegNet_V_3 have significantly improved (*p*-value < 0.0001). Specifically, compared with the mean DSC of the Baseline, this evaluation metric of the AC-RegNet_V1, AC-RegNet_V2, and AC-RegNet_V3 is improved by $-0.010 (\downarrow)$, $0.014 (\uparrow)$, and $0.074 (\uparrow)$, respectively. Besides, compared with the mean HD of the Baseline, this evaluation metric of the AC-RegNet_V1, AC-RegNet_V2, and AC-RegNet_V3 is improved by $4.568 (\downarrow)$, $-0.751 (\uparrow)$, and $- 8.693 (\uparrow)$, respectively. Meanwhile, compared with the mean ASSD of the Baseline, this evaluation metric of the AC-RegNet_V1, AC-RegNet_V1, AC-RegNet_V2, and AC-RegNet_V3 is improved by $-0.033 (\uparrow)$, $-0.987 (\uparrow)$, and $- 4.680 (\uparrow)$, respectively. Last, compared with the mean MSD of the Baseline, this evaluation metric of the AC-RegNet_V1, AC-RegNet_V2, and AC-RegNet_V3 is improved by $-362.148 (\uparrow)$, $-189.888 (\uparrow)$, and $- 519.776 (\uparrow)$, respectively.

Compared with the evaluation metrics of the AC-RegNet_V1, all these evaluation metrics of the AC-RegNet_V2 and AC-RegNet_V_3 have significantly improved (*p*-value < 0.0001). Specifically, compared with the mean DSC of the AC-RegNet_V1, this evaluation metric of the AC-RegNet_V2 and AC-RegNet_V3 is improved by 0.024 (\uparrow) and 0.084 (\uparrow), respectively. Besides, compared with the mean HD of the AC-RegNet_V1, this evaluation metric of the AC-RegNet_V2 and AC-RegNet_V2 and AC-RegNet_V2 and AC-RegNet_V3 is improved by -5.319 (\uparrow) and - 13.261 (\uparrow), respectively. Meanwhile, compared with the mean ASSD of the AC-RegNet_V1, this evaluation metric of the AC-RegNet_V2 and AC-RegNet_V1, this evaluation metric of the AC-RegNet_V2 and AC-RegNet_V3 is improved by -0.954 (\uparrow) and - 4.647 (\uparrow), respectively. Last, compared with the mean MSD of the AC-RegNet_V3 is improved by 172.26 (\downarrow) and - 157.628 (\uparrow), respectively.

Compared with the evaluation metrics of the AC-RegNet_V2, all these evaluation metrics of the AC-RegNet_V3 have significantly

improved (*p*-value < 0.0001). Specifically, compared with the mean DSC of the AC-RegNet_V2, this evaluation metric of the AC-RegNet_V3 is improved by 0.06 (\uparrow). Besides, compared with the mean HD of the AC-RegNet_V2, this evaluation metric of the AC-RegNet_V3 is improved by -7.942 (\uparrow). Meanwhile, compared with the mean ASSD of the AC-RegNet_V2, this evaluation metric of the AC-RegNet_V3 is improved by -3.693 (\uparrow). Last, compared with the mean MSD of the AC-RegNet_V2, this evaluation metric of the AC-RegNet_V3 is improved by -3.693 (\uparrow). Last, compared with the mean MSD of the AC-RegNet_V2, this evaluation metric of the AC-RegNet_V3 is improved by -3.29888 (\downarrow).

4.2 Comparing study results

Table 2 shows the descriptive statistics on the evaluation metrics of the 110 pairs of dynamic lung field registration images based on the comparing study. Meanwhile, Figure 7 shows the visual distribution of the evaluation metrics of the SimpleElastix and AC-RegNet_V3 based on the four pre-trained segmentation networks. Last, Figures 8–11 show the visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14–23) based on the SimpleElastixSegNet and AC-RegNet_V3 with the different pre-trained lung field segmentation networks.

First, the DSC (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the SimpleElastixSegNet with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 0.960 \pm 0.014, 0.959 \pm 0.020, 0.963 \pm 0.019, and 0.962 \pm 0.017, respectively. Meanwhile, the DSC (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the AC-RegNet_V3with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 0.991 \pm 0.018, 0.993 \pm 0.016, 0.993 \pm 0.017, and 0.993 \pm 0.016, respectively. Compared with the mean DSC of 0.961 [(0.96 + 0.959 + 0.963 + 0.962)/4] based on the SimpleElastixSegNet, this mean evaluation metric of 0.993 based on the AC-RegNet_V3 is improved by 0.0315 (\uparrow).

Second, the HD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the SimpleElastixSegNet with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 17.874 \pm 8.186, 19.147 \pm 9.136, 17.478 \pm 8.818, and 17.651 \pm 7.486, respectively. Meanwhile, the HD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the AC-RegNet_V3with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 12.512 \pm 11.338, 12.813 \pm 11.460, 12.449 \pm 11.450, and 13.661 \pm 11.760, respectively.

TABLE 2 The evaluation metrics of the comparing study.

Evaluation metrics	Network			110 pairs of images
	Segmentation	Lung field extraction	Registration	
DSC (mean ± SD*)	SegNet [25]	1	SimpleElastix [10]	0.960 (0.926-0.983) ± 0.014
	U-Net [26]	✓	SimpleElastix [10]	0.959 (0.926-0.983) ± 0.020
	ResU-Net ++ [17]	1	SimpleElastix [10]	0.963 (0.918-0.989) ± 0.019
	AttU-Net [18]	1	SimpleElastix [10]	0.962 (0.924-0.987) ± 0.017
	SegNet [25]	1	AC-RegNet_V3	0.991 (0.894–0.999) ± 0.018
	U-Net [26]	1	AC-RegNet_V3	0.993 (0.901-1.000) ± 0.016
	ResU-Net ++ [17]	1	AC-RegNet_V3	0.993 (0.898-1.000) ± 0.017
	AttU-Net [18]	1	AC-RegNet_V3	0.993 (0.905-1.000) ± 0.016
HD (mean ± SD*)	SegNet [25]	1	SimpleElastix [10]	17.874 (7.071-37.947) ± 8.186
	U-Net [26]	1	SimpleElastix [10]	19.147 (4.472-39.217) ± 9.136
	ResU-Net ++ [17]	1	SimpleElastix [10]	17.478 (4.000-36.674) ± 8.818
	AttU-Net [18]	1	SimpleElastix [10]	17.651 (6.403-37.577) ± 7.486
	SegNet [25]	1	AC-RegNet_V3	12.512 (1.000-42.579) ± 11.338
	U-Net [26]	1	AC-RegNet_V3	12.813 (1.000-42.202) ± 11.460
	ResU-Net ++ [17]	1	AC-RegNet_V3	12.449 (1.000-43.139) ± 11.450
	AttU-Net [18]	1	AC-RegNet_V3	13.661 (1.000-43.278) ± 11.760
ASSD (mean ± SD*)	SegNet [25]	1	SimpleElastix [10]	2.590 (1.115-4.795) ± 0.928
	U-Net [26]	1	SimpleElastix [10]	2.653 (0.762-5.812) ± 1.279
	ResU-Net ++ [17]	1	SimpleElastix [10]	2.506 (0.757-5.416) ± 1.278
	AttU-Net [18]	1	SimpleElastix [10]	2.525 (0.892-4.999) ± 1.126
	SegNet [25]	1	AC-RegNet_V3	0.654 (0.065-4.665) ± 1.025
	U-Net [26]	1	AC-RegNet_V3	0.550 (0.035-4.361) ± 0.911
	ResU-Net ++ [17]	1	AC-RegNet_V3	0.572 (0.025-4.825) ± 0.992
	AttU-Net [18]	1	AC-RegNet_V3	0.564 (0.019-4.702) ± 0.988
MSD (mean ± SD*)	SegNet [25]	1	SimpleElastix [10]	559.098 (156.463-
				1212.521) ± 225.013
	U-Net [26]	1	SimpleElastix [10]	577.797 (127.297– 1200.305) ± 269.190
	ResU-Net ++ [17]	1	SimpleElastix [10]	548.189 (116.112– 1183.068) ± 254.528
	AttU-Net [18]	1	SimpleElastix [10]	559.652 (131.187– 1117.780) ± 247.021
	SegNet [25]	1	AC-RegNet_V3	103.718 (13.923– 658.279) ± 116.340
	U-Net [26]	1	AC-RegNet_V3	92.878 (9.438–616.990) ± 106.752
	ResU-Net ++ [17]	✓	AC-RegNet_V3	94.750 (8.655-630.870) ± 94.750
	AttU-Net [18]	1	AC-RegNet_V3	92.028 (8.731–607.461) ± 109.771

* SD, standard deviation.

Compared with the mean HD of 18.038 based on the SimpleElastixSegNet, this mean evaluation metric of 12.859 based on the AC-RegNet_V3 is improved by -5.179 (\uparrow).

Third, the ASSD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the SimpleElastixSegNet with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 2.590 \pm 0.928,

 2.653 ± 1.279 , 2.506 ± 1.278 , and 2.525 ± 1.126 , respectively. Meanwhile, the ASSD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the AC-RegNet_V3with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 0.654 ± 1.025 , 0.550 ± 0.911 , 0.572 ± 0.992 , and 0.564 ± 0.988 , respectively. Compared with the mean ASSD of 2.569 based on the



SimpleElastixSegNet, this mean evaluation metric of 0.585 based on the AC-RegNet_V3 is improved by -1.984 (\uparrow).

Last, the MSD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the SimpleElastixSegNet with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 559.098 \pm 225.013, 577.797 \pm 269.190, 548.189 \pm 254.528, and 559.652 \pm 247.021, respectively. Meanwhile, the MSD (mean \pm SD) of these 110 pairs of dynamic lung field and registration images based on the AC-RegNet_ V3with the SegNet, U-Net, ResU-Net ++, and AttU-Net are 103.718 \pm 116.340, 92.878 \pm 106.752, 94.750 \pm 94.750, and 92.028 \pm 109.771, respectively. Compared with the mean MSD of 561.184 based on the SimpleElastixSegNet, this mean evaluation metric of 95.844 based on the AC-RegNet_V3 is improved by -465.341 (\uparrow).

Figures 8–11 show the visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14–23) based on the SimpleElastix and AC-RegNet_V3 with the different pre-trained lung field segmentation networks. The registration effect of the proposed AC-RegNet_V3 comprehensively surpasses that of the SimpleElastix. Visually, the source image (13) can be well aligned with target images (14–23) based on each lung field

segmentation network with AC-RegNet_V3. However, it is easy to observe a certain degree of tearing on the registration images (13_16, 13_22, and 13_23 marked by the red box) based on each lung field segmentation network with AC-RegNet_V3.

5 Discussion

This section conducts the following discussion and points out this study's limitations based on the experimental results.

5.1 Dynamic lung field registration driven by automatic segmentation technology

The automatic organ segmentation task based on CNN is the foundation and key to achieving quantitative analysis of regions of interest (ROI) in medical images, such as 3D chest CT images, magnetic resonance angiography images, and cardiac MRI images, etc. (Yang et al., 2024; Yang et al., 2024; Deng et al., 2024; Wang et al., 2024; Yang et al., 2024; Zeng et al., 2023; Duan et al., 2024; Deng et al., 2024; Den



FIGURE 8

Visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14-23) based on the SimpleElastixand AC-RegNet_V3 with the SegNet.



FIGURE 9

Visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14–23) based on the SimpleElastix and AC-RegNet_V3 with the Unet.



FIGURE 10

Visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14–23) based on the SimpleElastix and AC-RegNet_V3 with the ResU-Net ++.



FIGURE 11

Visual differences among the dynamic CXR, lung field, and registration images of the source image (13) and target images (14–23) based on the SimpleElastix and AC-RegNet_V3 with the AttU-Net.

2023). Similarly, this critical task also exists in the CXR images for ROI segmentation of the lung field.

Accurate lung field registration results may be more favored and exciting than the registration results of the entire CXR images for dynamic lung field analysis in clinical practice. The effective and accurate lung field segmentation of dynamic P-A CXR images will drive the development of dynamic lung field registration for clinical quantitative analysis. Meanwhile, the automatic lung field segmentation of the P-A CXR images is a data diversity problem, not a methodology problem (Yang et al., 2024). Therefore, our previous study trained several robust and standard segmentation networks of pathological lungs based on the CNN with cross-center static P-A CXR images and their diversity of disease. Besides, the data augmentation technology is applied to enrich the training set of the static P-A CXR images and relieve the engineering problem of generalization in these lung field segmentation networks (Yang et al., 2024; Chlap et al., 2021; Kiruthika and Khilar, 2024; Hasan and Abdulazeez, 2024). These robust and standard segmentation networks provide optional models for lung field segmentation in dynamic P-A chest CXR images. Based on the above, the proposal of these segmentation networks has laid a solid foundation for the lung field segmentation of cross-center dynamic P-A CXR images, allowing the subsequent registration of the dynamic lung fields. Actually, the registration evaluation metrics further indicate that these segmentation networks can meet the requirements for registering dynamic lung field images.

5.2 Dynamic lung field registration driven by static P-A CXR images

Compared with static P-A CXR images, dynamic CXR images have not been widely collected. Meanwhile, the lung undergoes a non-rigid and complex process of contraction and expansion during breathing (Santhanam, 2006). Compared to other static organs or tissues (such as the brain, bones, etc.), these non-rigid and complex deformations present significant challenges for lung field registration. Therefore, the limited number of dynamic CXR images and the non-rigid and complex deformations in the dynamic lung field make registering the dynamic lung fields challenging.

Based on the above, this paper fully utilizes static P-A CXR images to simulate the non-rigid and complex deformations in the dynamic lung field during the enhanced and final training of the basic architecture of AC-RegNet. Although AC-RegNet_V1 can produce certain deformations in dynamic images collected from the same person's breathing process, it lacks learning from dynamic CXR images during training, resulting in its inability to register dynamic CXR images well. When the enhanced and final training of the basic architecture of AC-RegNet, we are inspired by these lung field segmentation networks above and subsequently apply the affine transformation technology (Yang et al., 2024; Chlap et al., 2021), originally used for data augmentation for training segmentation networks, to simulate lung respiratory changes in the static P-A CXR and lung field images. Compared with AC-RegNet_V1, even if simulated dynamic images are added during the training process of AC-RegNet_2, the network needs to consider the registration of the entire CXR image, which imposes a great registration burden on the network. However, AC-RegNet_V3 eliminates this registration burden by focusing only on the lung fields. The significant improvement in registration evaluation metrics also benefits from the simulated changes in lung respiration mentioned above. Meanwhile, when the lung field changes too much during respiration, the lung field in the registration images may tear due to significant deformation of the source image for alignment with the target image. In addition, there is tearing in the lung field in the registered images, leading to a reasonably great value in HD of the evaluation metrics. However, we believe that this tearing is inevitable.

5.3 Providing the possibility for quantitative analysis of dynamic lung fields

The proposed fully automatic registration pipeline effectively registers dynamic lung field images to maintain lung morphology alignment during respiration. Thus, this may become a valuable tool for further quantitative analysis of dynamic lung fields, such as pulmonary airflow detection (Jiang et al., 2024) and air trapping location (Zhang et al., 2024).

Specifically, the most direct way is to use registered images from adjacent time points to detect the trajectory of lung field movement, thereby revealing the trajectory rules in the respiratory process. Besides, like the air trapping location based on expiratory and inspiratory CT images (Deng et al., 2024; Wang et al., 2024), a pair of lung field images with the maximum and minimum areas can be determined during a respiratory cycle. Then, the registration images based on this pair of lung field images can be used to locate the air trapping, which may provide a new avenue for diagnosing chronic obstructive pulmonary disease based on dynamic CXR images. Dynamic CXR images can also capture more images within one respiratory cycle compared to expiratory and inspiratory CT images at only 2-time points. Furthermore, the registration images based on adjacent CXR images can detect pulmonary airflow within the lung field (Jiang et al., 2024), which will undoubtedly help reveal pulmonary airflow movement patterns within a respiratory cycle.

5.4 Limitations

Although we propose a fully automatic three-stage registration pipeline based on static P-A CXR images that can effectively register dynamic lung field images to maintain lung field morphology alignment during respiration from an engineering perspective, our research still has certain limitations. We do not have sufficient dynamic CXR images to further validate the proposed registration pipeline's performance. Therefore, we encourage researchers to collect more dynamic CXR image images to validate the proposed registration pipeline's performance and subsequently perform the quantitative analysis of dynamic lung fields.

6 Conclusion

We propose a fully automatic three-stage registration pipeline for the dynamic lung field of CXR images based on AC-RegNet with the static P-A CXR images, which effectively addresses the issue of the inability to register dynamic lung field images for the quantitative analysis of lung fields. First, the dynamic lung field mask images are generated from a pre-trained standard lung field segmentation model with the dynamic CXR images. Then, a lung field abstraction model is designed to generate the dynamic lung field images based on the dynamic lung field mask images and their corresponding CXR images. Finally, we propose a three-step registration training method to train the AC-RegNet, obtaining the registration network of the dynamic lung field images. The results show that the mean evaluation metrics of registration images based on the four basic segmentation networks with AC-RegNet achieve the mean DSC of 0.991, 0.993, 0.993, and 0.993, mean HD of 12.512, 12.813, 12.449, and 13.661, mean ASSD of 0.654, 0.550, 0.572, and 0.564, and mean MSD of 559.098, 577.797, 548.189, and 559.652, respectively. Therefore, our proposed three-stage registration pipeline has demonstrated its effectiveness in dynamic lung field analysis in clinical practice.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

All patients provided written informed consent, and this study was approved by the Guangzhou Medical University Ethics Committee (Grant number: 2023-hg-ks-24, Approval Date: 28 August 2023). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YY: Formal analysis, Methodology, Software, Visualization, Writing – original draft. JZ: Formal analysis, Methodology, Software, Writing – original draft. PG: Formal analysis, Methodology, Software, Writing – original draft. QG: Formal analysis, Software, Validation, Writing – original draft. YG: Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing. ZC: Software, Writing – original draft. CL: Investigation, Validation, Writing – original draft. TW: Software, Writing – original draft. ZO: Investigation, Writing – original draft. HC: Data curation, Funding acquisition,

References

Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 2481–2495. doi: 10.1109/TPAMI.2016.2644615

Balakrishnan, G, Zhao, A, Sabuncu, M R, et al. An unsupervised learning model for deformable medical image registration. In Proceedings of the IEEE conference on computer vision and pattern recognition. (2018): 9252–9260. IEEE: Salt Lake City, UT

Chang, S., and Lee, S. C. (2021). A comparative study on the voxel values in alveolar bones acquired by MDCT and newly developed dental dual-energy CBCT. *Sensors* 21:7552. doi: 10.3390/s21227552

Chen, J., Zhong, Z., Wang, W., Yu, G., Zhang, T., and Wang, Z. (2022). Quantitative evaluation of diaphragmatic motion during forced breathing in chronic obstructive

Project administration, Resources, Supervision, Validation, Writing – review & editing. YK: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Zhongnanshan Medical Foundation of Guangdong Province, China (ZNSXS-20230001), the National Natural Science Foundation of China (62071311), the special program for key fields of colleges and universities in Guangdong Province (biomedicine and health) of China (2021ZDZX2008), China postdoctoral Foundation project (2024M750398), the National Key Research and Development Program of China (2022YFF0710800), the National Key Research and Development Program of China (2022YFF0710800), and the Heilongjiang Province Natural Science Foundation Joint Guidance Project (LH2024F044).

Acknowledgments

Thanks to the Department of Radiology, The Second Affiliated Hospital of Guangzhou Medical University, for providing the dataset.

Conflict of interest

YY, JZ, PG, TW, and ZO were employed by Shenzhen Lanmage Medical Technology Co., Ltd. QG was employed by Neusoft Medical System Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

pulmonary disease patients using dynamic chest radiography. *Front. Integr. Neurosci.* 16:842404. doi: 10.3389/fnint.2022.842404

Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., and Haworth, A. (2021). A review of medical image data augmentation techniques for deep learning applications. *J. Med. Imaging Radiat. Oncol.* 65, 545–563. doi: 10.1111/1754-9485.13261

Deng, X., Li, W., Yang, Y., Wang, S., Zeng, N., Xu, J., et al. (2024). COPD stage detection: leveraging the auto-metric graph neural network with inspiratory and expiratory chest CT images. *Med. Biol. Eng. Comput.* 62, 1733–1749. doi: 10.1007/s11517-024-03016-z

Duan, W., Guo, Y., Yang, Y., Zeng, N., Wang, S., Xu, J., et al. (2023). Exploring the effectiveness of multi-objective training for organ substructure segmentation in medical imaging. *Stud. Health Technol. Inform.* 308, 146–154. doi: 10.3233/SHTI230835

Ferrante, E, Oktay, O, Glocker, B, and Milone, D. H. On the adaptability of unsupervised CNN-based deformable image registration to unseen image domains. Machine learning in medical imaging: 9th international workshop, Granada, Spain: Springer. (2018) 294–302.

Guo, Y., Yang, Y., Cao, F., Li, W., Wang, M., Luo, Y., et al. (2022). Novel survival features generated by clinical text information and radiomics features may improve the prediction of ischemic stroke outcome. *Diagnostics* 12:1664. doi: 10.3390/diagnostics12071664

Guo, Y., Yang, Y., Cao, F., Liu, Y., Li, W., Yang, C., et al. (2022). Radiomics features of DSC-PWI in time dimension may provide a new chance to identify ischemic stroke. *Front. Neurol.* 13:889090. doi: 10.3389/fneur.2022.889090

Guo, Y., Yang, Y., Cao, F., Wang, M., Luo, Y., Guo, J., et al. (2022). A focus on the role of DSC-PWI dynamic radiomics features in diagnosis and outcome prediction of ischemic stroke. J. Clin. Med. 11:5364. doi: 10.3390/jcm11185364

Guo, Y., Yang, Y., Wang, M., Luo, Y., Guo, J., Cao, F., et al. (2022). The combination of whole-brain features and local-lesion features in DSC-PWI may improve ischemic stroke outcome prediction. *Life* 12:1847. doi: 10.3390/life12111847

Hasan, D., and Abdulazeez, A. M. (2024). Lung segmentation from chest X-ray images using Deeplabv3plus-based CNN model. *Indones. J. Comput. Sci.* 13, 13–24. doi: 10.33022/ijcs.v13i1.3700

Howell, J. D. (2016). Early clinical use of the X-ray. Trans. Am. Clin. Climatol. Assoc. 127, 341-349

Iqbal, I., Ullah, I., Peng, T., Wang, W., and Ma, N. (2025). An end-to-end deep convolutional neural network-based data-driven fusion framework for identification of human induced pluripotent stem cell-derived endothelial cells in photomicrographs. *Eng. Appl. Artif. Intell.* 139:109573. doi: 10.1016/j.engappai.2024.109573

Iqbal, I., Younus, M., Walayat, K., Kakar, M. U., and Ma, J. (2021). Automated multiclass classification of skin lesions through deep convolutional neural network with dermoscopic images. *Comput. Med. Imaging Graph.* 88:101843. doi: 10.1016/j.compmedimag.2020.101843

Jaderberg, M., Simonyan, K., and Zisserman, A. (2015). Spatial transformer networks. *Adv. Neural Inf. Process. Syst.* 28.

Jafar, A., Hameed, M. T., Akram, N., Waqas, U., Kim, H. S., and Naqvi, R. A. (2022). CardioNet: automatic segmentation to calculate the cardiothoracic ratio for cardiomegaly and other chest diseases. *J. Pers. Med.* 12:988. doi: 10.3390/jpm12060988

Jha, D, Smedsrud, P H, Riegler, M A, et al. (2019). ResUNet++: An advanced architecture for medical image segmentation. In: Proceedings of the 2019 IEEE International Symposium on Multimedia ISM 2019, IEEE: Laguna Hills, CA. 225–2255.

Jiang, F., Hirano, T., Liang, C., Zhang, G., Matsunaga, K., and Chen, X. (2024). Multiscale simulations of pulmonary airflow based on a coupled 3D-1D-0D model. *Comput. Biol. Med.* 171:108150. doi: 10.1016/j.compbiomed.2024.108150

Kiruthika, K., and Khilar, R. (2024). Segmentation of lung on P-A CXR images based on CXR-auto encoder segmentation with MRF. *Opt. Quant. Electron.* 56, 1–15. doi: 10.1007/s11082-024-06507-3

Kuang, D. Cycle-consistent training for reducing negative jacobian determinant in deep registration networks. In Simulation and Synthesis in Medical Imaging: 4th International Workshop. Springer-Verlag: Berlin, Heidelberg. (2019).

Long, J, Shelhamer, E, and Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. (2015). Boston, MA: IEEE. 3431–3440.

Mansilla, L., Milone, D. H., and Ferrante, E. (2020). Learning deformable registration of medical images with anatomical constraints. *Neural Netw.* 124, 269–279. doi: 10.1016/j.neunet.2020.01.023

Marstal, K, Berendsen, F, Staring, M, et al. SimpleElastix: a user-friendly, multi-lingual library for medical image registration. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops. (2016) Las Vegas, NV: IEEE. 134–142.

Nie, Q., Zhang, X., Hu, Y., Gong, M., and Liu, J. (2024). Medical image registration and its application in retinal images: a review. *Vis. Comput. Ind. Biomed. Art.* 7:21.

Rahunathan, S, Stredney, D, Schmalbrock, P, et al. Image registration using rigid registration and maximization of mutual information. In Proceedings of the 13th annual medicine meets virtual reality conference. (2005). Long Beach, CA: MMVR13

Ronneberger, O, Fischer, P, and Brox, T. (2015) U-net: Convolutional Networks for Biomedical Image Segmentation. In: International conference on medical image computing and computer-assisted intervention, Springer, Cham.

Santhanam, A. (2006). Modeling, simulation, and visualization of 3d lung dynamics. Orlando, FL: University of Central Florida.

Seah, J. C. Y., Tang, C. H. M., Buchlak, Q. D., Holt, X. G., Wardman, J. B., Aimoldin, A., et al. (2021). Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: a retrospective, multireader multicase study. *Lancet Digit. Health* 3, e496–e506. doi: 10.1016/S2589-7500(21)00106-0

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. and Manzagol, P-A. (2010). Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* 11, 3371–3408. Available at: https://dl.acm.org/doi/10.5555/1756006.1953039

Wang, S., Li, W., Zeng, N., Xu, J., Yang, Y., Deng, X., et al. (2024). Acute exacerbation prediction of COPD based on auto-metric graph neural network with inspiratory and expiratory chest CT images. *Heliyon* 10:e28724. doi: 10.1016/j.heliyon.2024.e28724

Wang, S, Li, L, and Zhuang, X. AttU-net: Attention U-net for brain tumor segmentation. In International MICCAI Brainlesion workshop. Cham: Springer International Publishing, (2021). 302–311.

Xiao, H., Xue, X., Zhu, M., Jiang, X., Xia, Q., Chen, K., et al. (2023). Deep learningbased lung image registration: a review. *Comput. Biol. Med.* 165:107434. doi: 10.1016/j.compbiomed.2023.107434

Yang, Y., and Guo, Y. (2024). Ischemic stroke outcome prediction with diversity features from whole brain tissue using deep learning network. *Front. Neurol.* 15:1394879. doi: 10.3389/fneur.2024.1394879

Yang, Y, Guo, Y, Guo, J, et al. A method of abstracting single pulmonary lobe from computed tomography pulmonary images for locating COPD. In Proceedings of the fourth international conference on biological information and biomedical engineering. (2020): 1–6. Association for Computing Machinery: China

Yang, G., Huang, J., He, Y., Chen, Y., Wang, T., Jin, C., et al. (2022). GCP-net: a gating context-aware pooling network for cervical cell nuclei segmentation. *Mob. Inf. Syst.* 2022, 1–14. doi: 10.1155/2022/7511905

Yang, Y., Li, Q., Guo, Y., Liu, Y., Li, X., Guo, J., et al. (2021). Lung parenchyma parameters measure of rats from pulmonary window computed tomography images based on ResU-net model for medical respiratory researches. *Math. Biosci. Eng.* 18, 4193–4211. doi: 10.3934/mbe.2021210

Yang, Y., Li, W., Guo, Y., Zeng, N., Wang, S., Chen, Z., et al. (2022). Lung radiomics features for characterizing and classifying COPD stage based on feature combination strategy and multi-layer perceptron classifier. *Math. Biosci. Eng.* 19, 7826–7855. doi: 10.3934/mbe.2022366

Yang, Y., Zheng, J., Guo, P., Wu, T., Gao, Q., Guo, Y., et al. (2024). Automatic cardiothoracic ratio calculation based on lung fields abstracted from chest X-ray images without heart segmentation. *Front. Physiol.* 15:6912. doi: 10.3389/fphys.2024.1416912

Yang, Y., Zheng, J., Guo, P., Wu, T., Gao, Q., Zeng, X., et al. (2024). Hemi-diaphragm detection of chest X-ray images based on convolutional neural network and graphics. *J. Xray Sci. Technol.* 32, 1273–1295. doi: 10.3233/XST-240108

Yeap, P. L., Wong, Y. M., Ong, A. L. K., Tuan, J. K. L., Pang, E. P. P., Park, S. Y., et al. (2023). Predicting dice similarity coefficient of deformably registered contours using Siamese neural network. *Phys. Med. Biol.* 68:155016. doi: 10.1088/1361-6560/ace6f0

Zaman, A., Hassan, H., Zeng, X., Khan, R., Lu, J., Yang, H., et al. (2024). Adaptive feature medical segmentation network: an adaptable deep learning paradigm for high-performance 3D brain lesion segmentation in medical imaging. *Front. Neurosci.* 18:1363930. doi: 10.3389/fnins.2024.1363930

Zeng, X., Guo, Y., Zaman, A., Hassan, H., Lu, J., Xu, J., et al. (2023). Tubular structure segmentation via multi-scale reverse attention sparse convolution. *Diagnostics* 13:2161. doi: 10.3390/diagnostics13132161

Zhang, Y., Tanabe, N., Sato, S., Shiraishi, Y., Maetani, T., Sakamoto, R., et al. (2024). Longitudinal changes in respiratory reactance in patients with COPD: associations with longitudinal change in air-trap**, exacerbations, and mortality. *Respir. Physiol. Neurobiol.* 322:104216. doi: 10.1016/j.resp.2024.104216