




Fully Automatic Registration Methods for Chest X-Ray Images

Yu-Ching Lee¹ · Muhammad Adil Khalil² · Jui-Huan Lee² · Abdan Syakura¹ · Yi-Fang Ding³ · Ching-Wei Wang^{1,2} 

Received: 3 November 2020 / Accepted: 14 October 2021 / Published online: 3 November 2021
© Taiwanese Society of Biomedical Engineering 2021

Abstract

Purpose Image registration is important in medical applications accomplished by improving healthcare technology in recent years. Various studies have been proposed in medical applications, including clinical track of events and updating the treatment plan for radiotherapy and surgery. This study presents a fully automatic registration system for chest X-ray images to generate fusion results for difference analysis. Using the accurate alignment of the proposed system, the fusion result indicates the differences in the thoracic area during the treatment process.

Methods The proposed method consists of a data normalization method, a hybrid L-SVM model to detect lungs, ribs and clavicles for object recognition, a landmark matching algorithm, two-stage transformation approaches and a fusion method for difference analysis to highlight the differences in the thoracic area. In evaluation, a preliminary test was performed to compare three transformation models, with a full evaluation process to compare the proposed method with two existing elastic registration methods.

Results The results show that the proposed method produces significantly better results than two benchmark methods (P -value ≤ 0.001). The proposed system achieves the lowest mean registration error distance (MRED) (8.99 mm, 23.55 pixel) and the lowest mean registration error ratio (MRER) w.r.t. the length of image diagonal (1.61%) compared to the two benchmark approaches with MRED (15.64 mm, 40.97 pixel) and (180.5 mm, 472.69 pixel) and MRER (2.81%) and (32.51%), respectively.

Conclusions The experimental results show that the proposed method is capable of accurately aligning the chest X-ray images acquired at different times, assisting doctors to trace individual health status, evaluate treatment effectiveness and monitor patient recovery progress for thoracic diseases.

Keywords Fully automatic image registration · Chest X-ray image comparison · Difference analysis

1 Introduction

With the progression of healthcare technology, radiological images are increasingly used in medical research, diagnosis, treatment planning and basic science [1]. A variety of studies have been proposed in recent years for the registration

of radiological images in various medical applications, including image-guided surgical systems [2, 3] and radiotherapy [4, 5]. Registration is the process to find the optimal transformation matrix from the coordinate of one data to another of the same scene taken at different times or different modalities, so that all features emerging in both images are overlapped and aligned. Image registration is generally achieved through four basic steps, including image rescaling, feature detection, feature matching, and transformation function construction [6].

An automatic registration system could be classified as a semi-automatic and fully automatic approach. Semi-automatic systems require corresponding landmarks annotated by medical experts for both of the images. Several studies have been proposed for developing a semi-automatic registration system, which had been applied to several medical applications [7–9]. However, manually annotated landmarks

✉ Ching-Wei Wang
cweiwang@mail.ntust.edu.tw

¹ Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei City, Taiwan

² Graduate Institute of Biomedical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan

³ Department of Otolaryngology, Wan Fang Hospital, Taipei Medical University, Taipei City, Taiwan

are time-consuming, subjective and error-prone due to fatigue or image blurriness [10, 11]. Moreover, registration of the chest X-ray images is a challenging task due to variations in data appearance, imaging artefacts and complex data deformation problems, making existing registration approaches unstable and performing poorly. Therefore, we present a fully automatic registration system in this study. The proposed method automatically detects features, extracts corresponding landmarks and produces an optimal transformation function without manual intervention. Geometric transformation can be classified into rigid and non-rigid transformations. Rigid transformation is conducted by linear transformation with translation, rotation, scaling and shearing parameters [12, 13]. Considering that the geometric difference may be neglected in the rigid transformation [14], non-rigid transformation aims to warp local geometric features [15, 16], which allows the crooked structural deformation to fix the geometric difference in the deformation.

In this study, we implement rigid and non-rigid transformation in the system to combine the advantages of both types of approaches. The purpose of this study is to develop and validate a fully automatic registration system to accurately align the chest X-ray images acquired at different times during the treatment. Consequently, the fused result could be used for the difference analysis for an effective treatment.

2 Methods

2.1 Methods Overview

In brief, the proposed method is constructed with data preprocessing to normalize input and training datasets, a hybrid L-SVM model to detect lungs, ribs and clavicles for object recognition, an Absolute Distance Matching Algorithm (ADMA) to identify and match corresponding landmarks of input images, two-stage transformation approaches, and difference analysis to highlight the differences in the thoracic area between the two images. The flow diagram of the proposed system is shown in Fig. 1.

2.2 Data Normalization

The data normalization process includes histogram matching [17, 18] and scaling standardization. Histogram matching in this study is an approach to modify the histogram distribution of input images corresponding to a pre-determined reference image for the luminance compensation and contrast enhancement. Figure 2 presents the illustration of histogram matching and the reference image is selected based on the image contrast quality (see Fig. 2a).

The first step of data normalization is to compute the histogram of input h_i and reference images h_r .

$$h[v] = \frac{1}{w \times h} \sum_{m=0}^{h-1} \sum_{n=0}^{w-1} \sigma[v, y_r[m, n]] \quad (1)$$

$$\sigma[a, b] = \begin{cases} 1, & \text{if } a = b. \\ 0, & \text{otherwise.} \end{cases}$$

where w and h denote the width and height of image, $y_r[m, n]$ denotes the image intensity of the reference image. Then, the cumulative histogram of reference H_r and input images H_i are computed.

$$H_r[j] = \sum_{i=0}^j h_r[i] \quad (2)$$

$$H_i[j] = \sum_{i=0}^j h_i[i] \quad (3)$$

Based on the difference of cumulative histogram H_d of H_r and H_i in Eq. 4, system finds an output level R for each input level T , and matches $H_i[T]$ to $H_r[R]$ by lookup entry in Eq. 5.

$$|H_r[R] - H_i[T]| = \min_z |H_r[R] - H_i[z]| \quad (4)$$

$$lookup[R] = T \quad (5)$$

Then, scaling standardization is conducted to resize the width of images to 512 pixels and the height of images to the same ratio with width.

2.3 Hybrid L-SVM Model

A hybrid L-SVM model is composed of six L-SVM models for detection of left and right lungs, ribs, clavicle and middle clavicle. We built the hybrid L-SVM model based on Felzenszwalb Histograms of Oriented Gradients (FHOG) features [19, 20] and Linear Support Vector Machines (L-SVM) model [21]. Figure 3 presents the illustration from training templates to the hybrid L-SVM model. To obtain 31-dimensional FHOG features for shape description, nine contrast insensitive gradient orientations, four dimensions capturing overall gradient magnitude, and 18 contrast sensitive features are computed from training templates.

To compute the gradient magnitude of training template, the gradients of horizontal and vertical approximations (G_h and G_v) are defined as follows.

$$G_h = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} * I(x, y) \quad (6)$$

$$G_v = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} * I(x, y) \tag{7}$$

Then, the gradient magnitude $G(x, y)$ and the angle of orientation $\theta(x, y)$ are further defined in Eqs. 8 and 9.

$$G(x, y) = \sqrt{(G_h)^2 + (G_v)^2} \tag{8}$$

$$\theta(x, y) = \text{atan}\left(\frac{G_h}{G_v}\right) \tag{9}$$

In order to reduce the dimension of feature vector, the definition proposed by Felzenszwalb [19] is utilized in Eqs. 10 and 11. When $k > 8$, 13-dimensional feature is equal to $a_0, a_1, \dots, a_8 \cup b_0, b_1, \dots, b_3$.

$$a_k(i, j) = \begin{cases} 1, & \text{if } j = k. \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

$$b_k(i, j) = \begin{cases} 1, & \text{if } i = k. \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

where a_k is computed by summing the four normalizations for nine orientations, and b_k is computed by summing nine orientations for four normalizations from output feature.

Based on the FHOg descriptor extracted from the training database, each L-SVM model finds a hyperplane for classification in the optimal layer of the multiple scaling feature pyramid. To detect objects in the testing images, the L-SVM model scans across all scales and positions of input images and computes an overall score for each root location according to the best possible placement of the parts.

$$\text{Score}(p_0) = \max_{P_0, \dots, P_n} (P_0, \dots, P_n) \tag{12}$$

The location yielding a high-scoring result defines the success detection of the target object.

2.4 Spin Assisted Algorithm (SAA)

After the left and right lung detection with the hybrid L-SVM model, we developed an algorithm called Spin Assisted Algorithm (SAA). This was utilized to calibrate a crooked body to the correct position to increase the detection accuracy of the left and right ribs, clavicle and middle clavicle.

Using the line connected with the top coordinates of the left and right lung as a baseline L_b , SAA computes the angle between the center line of chest L_c and the normal geometry of the baseline L_n for rotation. Figure 4 presents the illustration of SAA. In SAA, P_c , P_h and P_v are separately defined in Eqs. 13, 14 and 15. P_c is the central point of L_b , P_h is the

intersection of the L_n and the highest horizontal line of chest X-ray image, and P_v is the intersection of L_n and L_c .

$$x_{P_c} = \frac{x_{P_l} + x_{P_r}}{2} \quad \text{and} \quad y_{P_c} = \frac{y_{P_l} + y_{P_r}}{2} \tag{13}$$

$$x_{P_h} = -1 \div \frac{x_{P_l} - x_{P_r}}{y_{P_l} - y_{P_r}} \times (y_{P_h} - y_{P_c}) + x_{P_c}, y_{P_h} = 0 \tag{14}$$

$$x_{P_v} = 256, y_{P_v} = (x_{P_v} - x_{P_h}) \times -\left(\frac{x_{P_l} - x_{P_r}}{y_{P_l} - y_{P_r}}\right) + y_{P_h} \tag{15}$$

The rotation angle θ of the chest X-ray image is computed from the angle of L_n and L_c .

$$\theta = \text{atan}\left(\frac{x_{P_v} - x_{P_h}}{y_{P_v} - y_{P_h}}\right) \tag{16}$$

The central point of the input image P_i and rotate equation are defined in Eqs. 17, 18 and 19.

$$x_{P_i} = \frac{I'_{width}}{2} \quad \text{and} \quad y_{P_i} = \frac{I'_{height}}{2} \tag{17}$$

$$x' = (x - x_{P_i}) \times \cos(\theta) - (y - y_{P_i}) \times \sin(\theta) + x_{P_i} \tag{18}$$

$$y' = (x - x_{P_i}) \times \sin(\theta) - (y - y_{P_i}) \times \cos(\theta) + y_{P_i} \tag{19}$$

2.5 Absolute Distance Matching Algorithm (ADMA)

To precisely match corresponding landmarks, we built an algorithm called Absolute Distance Matching Algorithm (ADMA). To match corresponding landmarks $(\mathbf{P}_i^T, \mathbf{P}_i^S)$, ADMA selects the central points of the detected clavicle (p_{cc}^T, p_{cc}^S) and top points of the left and right lung of target and source images $(p_{ll}^T, p_{rl}^T, p_{ll}^S, p_{rl}^S)$ for the base points $(p_{lb}^T, p_{rb}^T, p_{lb}^S, p_{rb}^S)$ definition. Then, ADMA calculates the vertical distance $(A_l^T, A_r^T, A_l^S, A_r^S)$ as a matching value from the central point of the detected ribs $(\mathbf{P}_{lr}^T, \mathbf{P}_{rr}^T, \mathbf{P}_{lr}^S, \mathbf{P}_{rr}^S)$ to base points and sets up thresholds (t_l and t_r) with an average height of detected ribs. ADMA selects matching landmarks when the subtraction of matching values is lower than the threshold. After landmarks matching, ADMA transforms the landmarks to their original position before data preprocessing. Figure 5 illustrates ADMA.

Considering the transformation accuracy, ADMA selects at most seven corresponding landmarks of the target image I_t and source image I_s , which are defined as point array \mathbf{P}_i^T and \mathbf{P}_i^S . Based on different detection condition occurring in different parts of the chest X-ray image, we set up 12 flags $(F_{ll}^T, F_{ll}^S, F_{rl}^T, F_{rl}^S, F_{ll}^T, F_{ll}^S, F_{rl}^T, F_{rl}^S, F_{lr}^T, F_{lr}^S, F_{rr}^T, F_{rr}^S, F_c^T, F_c^S, F_{mc}^T, F_{mc}^S)$ to reflect

the result whether the hybrid L-SVM model detects feature in the regions of the left lung (F_{ll}^T) in I_t , left lung (F_{ll}^S) in I_s , right lung (F_{rl}^T) in I_t , right lung (F_{rl}^S) in I_s , left ribs (F_{lr}^T) in I_t , left ribs (F_{lr}^S) in I_s , right ribs (F_{rr}^T) in I_t , right ribs (F_{rr}^S) in I_s , clavicle (F_c^T) in I_t , clavicle (F_c^S) in I_s , middle clavicle (F_{mc}^T) in I_t and middle clavicle (F_{mc}^S) in I_s .

$$12flags = \begin{cases} 1, & \text{detected object of area} > 0. \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

After SAA, the top points of the left and right lungs (p_{ll}^T for left lung of target image, p_{rl}^T for right lung for target image, p_{ll}^S for left lung of source image, p_{rl}^S for right lung of source image) are located and utilized for base points definition. Four basepoints of left and right, matching the source and target images (p_{lb}^T for left matching of target image, p_{lb}^S for left matching of target image, p_{rb}^T for right matching of target image and p_{rb}^S for right matching of source image) are defined in Eqs. 22, 23, 25 and 26. The basepoints are affected by the detection condition of the lungs, ribs and clavicle, and the central point of the clavicle (p_{cc}^T for target image, p_{cc}^S for source image) is defined by the central point of the clavicle (p_{cl}^T for target image, p_{cl}^S for source image) and middle clavicle (p_{mc}^T for target image, p_{mc}^S for source image) detected by the hybrid L-SVM model in Eqs. 21 and 24.

In I_t :

$$P_{cc}^T = \begin{cases} P_{mc}^T, & F_{mc}^T > 0. \\ P_{cl}^T, & F_{cl}^T > 0. \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

$$P_{lb}^T = \begin{cases} P_{ll}^T, & F_{ll}^T > 0. \\ P_{cc}^T, & F_{cc}^T > 0 \text{ or } F_{mc}^T > 0. \\ P_{rl}^T, & F_{rl}^T > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

$$P_{rb}^T = \begin{cases} P_{rl}^T, & F_{rl}^T > 0. \\ P_{cc}^T, & F_{cc}^T > 0 \text{ or } F_{mc}^T > 0. \\ P_{ll}^T, & F_{ll}^T > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

In I_s :

$$P_{cc}^S = \begin{cases} P_{mc}^S, & F_{mc}^S > 0. \\ P_{cl}^S, & F_{cl}^S > 0. \\ 0, & \text{otherwise.} \end{cases} \quad (24)$$

$$P_{lb}^S = \begin{cases} P_{ll}^S, & F_{ll}^S > 0. \\ P_{cc}^S, & F_{cc}^S > 0 \text{ or } F_{mc}^S > 0. \\ P_{rl}^S, & F_{rl}^S > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

$$P_{rb}^S = \begin{cases} P_{rl}^S, & F_{rl}^S > 0. \\ P_{cc}^S, & F_{cc}^S > 0 \text{ or } F_{mc}^S > 0. \\ P_{ll}^S, & F_{ll}^S > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

To match the corresponding landmarks for ribs, four point arrays ($\mathbf{P}_{lr}^T, \mathbf{P}_{lr}^S, \mathbf{P}_{rr}^T, \mathbf{P}_{rr}^S$) are separately defined as the central points of left detected ribs of the target image (\mathbf{P}_{lr}^T), left detected ribs of source image (\mathbf{P}_{lr}^S), right detected ribs of target image (\mathbf{P}_{rr}^T) and of right detected ribs of source image (\mathbf{P}_{rr}^S). Based on the vertical distance from the central point of every detected rib to the basepoint, four arrays ($A_{lr}^T, A_{lr}^S, A_{rr}^T, A_{rr}^S$) are generated. A_{lr}^T, A_{lr}^S and A_{rr}^T, A_{rr}^S are defined as the vertical distance from \mathbf{P}_{lr}^T to $p_{lb}^T, \mathbf{P}_{lr}^S$ to $p_{lb}^S, \mathbf{P}_{rr}^T$ to $p_{rb}^T, \mathbf{P}_{rr}^S$ to p_{rb}^S . In addition, four thresholds ($t_l^T, t_l^S, t_r^T, t_r^S$) are generated as matching thresholds. t_l^T, t_l^S, t_r^T , and t_r^S are separately defined as the average height of the detected left ribs in the target image, left ribs in the source image, right ribs in the target image, and right ribs in the source image. The corresponding landmark matching of left and right ribs are conducted to compare whether the subtraction of A_{lr}^S and A_{lr}^T is lower than the left threshold t_l Eq. 27 and whether the subtraction of A_{rr}^S and A_{rr}^T is lower than the right threshold t_r Eq. 28. \mathbf{P}_i^T and \mathbf{P}_i^S are further defined in Eqs. 29 and 30.

$$t_l = (t_l^T + t_l^S) \div 2 \times \delta, \quad (27)$$

$$t_r = (t_r^T + t_r^S) \div 2 \times \delta, \quad (28)$$

where δ and δ are empirically determined; $\delta = 0.8, \delta = 0.8$.

In I_t :

$$P_i^T = \begin{cases} P_{ll}^T, & F_{ll}^T > 0 \text{ and } F_{ll}^S > 0. \\ P_{cc}^T, & (F_c^T > 0 \text{ or } F_{mc}^T > 0) \text{ and} \\ & (F_c^S > 0 \text{ or } F_{mc}^S > 0). \\ P_{rl}^T, & F_{rl}^T > 0 \text{ and } F_{rl}^S > 0. \\ \mathbf{P}_{lr}^T, & i < 5 \text{ and } \varphi < t_l. \\ \mathbf{P}_{rr}^T, & i < 7 \text{ and } \tau < t_r. \end{cases} \quad (29)$$

$$\mathbf{P}_i^T = [P_0^T, P_1^T, \dots, P_i^T]$$

where $\tau = \text{abs}(A_{lj}^S - A_{lk}^T)$ and $\varphi = \text{abs}(A_{rm}^S - A_{rn}^T)$.

In I_s :

$$P_i^S = \begin{cases} P_{ll}^S, & F_{ll}^S > 0 \text{ and } F_{ll}^T > 0. \\ P_{cc}^S, & (F_c^S > 0 \text{ or } F_{mc}^S > 0) \text{ and} \\ & (F_c^T > 0 \text{ or } F_{mc}^T > 0). \\ P_{rl}^S, & F_{rl}^S > 0 \text{ and } F_{rl}^T > 0. \\ \mathbf{P}_{lr}^S, & i < 5 \text{ and } \tau < t_l. \\ \mathbf{P}_{rr}^S, & i < 7 \text{ and } \varphi < t_r. \end{cases} \quad (30)$$

$$\mathbf{P}_i^S = [P_0^S, P_1^S, \dots, P_i^S]$$

where $\tau = \text{abs}(A_{lj}^S - A_{lk}^T)$ and $\varphi = \text{abs}(A_{rm}^S - A_{rn}^T)$.

After ADMA localizes transformation landmarks, we compute corresponding location of transformation landmarks in original chest X-ray images. The revert process follows the transformation, aiming at original images to decrease the image distortion in the corresponding landmarks extraction. To transform landmark sets matched by ADMA correspond to original images before data pre-processing, the revert angle ϑ is defined as $-\theta$ Eq. 16. x_p and y_p are defined as the coordinates of landmark sets after revert-rotation. x'_p and y'_p are defined as the coordinates of landmark sets transformed to the original scaling of input images.

$$\begin{aligned} x'_p &= (I_{width} - 1) - \left(\frac{I_{width} - 1}{I'_{width} - 1} \times ((I'_{width} - 1) - x_{P_i}) \right) \\ y'_p &= (I_{height} - 1) - \left(\frac{I_{height} - 1}{I'_{height} - 1} \times ((I'_{height} - 1) - y_{P_i}) \right) \end{aligned} \quad (31)$$

2.6 Singular Value Decomposition (SVD)

After ADMA, we adopted a two-stage registration, including global and elastic registration approaches for transformation. In the previous research, global registration is regarded as an Absolute Orientation Problem (AOP) [22]. It finds the optimal transformation function consisting of rotation, translation, and scaling for the source image. Based on the corresponding landmark sets extracted from input images, global registration calculates the optimal rigid transformation matrix for the source image to obtain the minimum error registration. The global registration is defined by finding the optimal transformation (rotation R, translation T and scaling S) from a corresponding landmark set T_i of the target image to the corresponding landmark set S_i of the source image, where $i \in [1, n]$ and n are defined as the number of corresponding landmarks matched by ADMA. The corresponding landmarks of S_i after global registration is defined.

$$S_i = (SRT_i + T) + N_i \quad (32)$$

where N_i is defined as a noise vector. To find the optimal rigid transformation matrix, the minimum error distance ζ is denoted in Eq. 33.

$$\zeta = \frac{1}{n} \sum_{i=1}^n (S_i - (SRT_i + T))^2 \quad (33)$$

The approach of global registration of the proposed method is called Singular Value Decomposition (SVD), proposed by Umeyama [23]. The transformation matrix of SVD (M_s) is defined.

$$M_s = \sum_{i=1}^n T'_i (S'_i)^T \quad (34)$$

where $T'_i = T_i - \bar{T}_i$ and $(S'_i)^T = (S_i - \bar{S}_i)^T$. The SVD matrix (M_s) is composed of right-singular orthogonal matrix R_s , left-singular orthogonal matrix L_s , and non-zero diagonal singular matrix α .

$$M_s = L_s \alpha R_s^T \quad (35)$$

L_s and R_s are defined in 36 by multiplying with M_s^T

$$\begin{aligned} M_s M_s^T L_s &= L_s \alpha^2 \\ M_s^T M_s R_s &= R_s \alpha^2 \end{aligned} \quad (36)$$

After three components of SVD matrix M_s are determined, the optimal transformation (rotation R, translation T and scaling S) are computed in Eqs. 37, 38 and 39.

$$\begin{aligned} R &= M_s \omega L_s^T \\ \omega &= \begin{cases} I, & \text{if } \det(R) \geq 0. \\ \det(1, 1, -1), & \text{if } \det(R) < 0. \end{cases} \end{aligned} \quad (37)$$

where ω is defined as identical matrix I.

$$T = \bar{S}_i - (R \times T_i) \quad (38)$$

$$S = \frac{\alpha}{\sum_{i=1}^n \det(T'_i)} \quad (39)$$

The source image after alignment I'_s is defined.

$$I'_s = SR \times I_s + T \quad (40)$$

2.7 Elastic Registration

In order to ensure the correct alignment result, we further implement elastic registration called Elastix [24, 25] after global registration. The affine and B-spline elastic registration methods were conducted sequentially, and the parameters of individual methods were optimized separately. Figure 6 presents the process of elastic transformation in the proposed system.

The intensity-based registration is formulated as an optimization problem, with the cost function \mathcal{C} minimized by the geometric transformation T for image registration. \mathcal{C} is measured by the similarity between the target image I_t and source image I_s after SVD.

$$\hat{T}_\mu = \arg \min_{T_\mu} \mathcal{C}(\mu; I_t, I_s) \quad (41)$$

where μ indicates the transformation has been parameterized. The vector μ contains the values of the transformation parameters.

In the elastic registration, a Gaussian pyramid is conducted to the target and source image after SVD to create image pyramids by down sampling and smoothing. The Gaussian function $G(\sigma_r)$ is defined to reduce the data and transformation complexity of registration.

$$G(\sigma_r) = \frac{1}{\sqrt{2\pi\sigma_r^x\sigma_r^y\sigma_r^z}} \hat{e}^{-\left(\frac{x^2}{2(\sigma_r^x)^2} + \frac{y^2}{2(\sigma_r^y)^2} + \frac{z^2}{2(\sigma_r^z)^2}\right)} \tag{42}$$

The input target and source images are separately defined as $I_t(x) : \forall_t \subset \hat{S} \rightarrow S$ and $I_s(x) : \forall_s \subset \hat{S} \rightarrow S$, where $x \in \hat{S}$ represents the coordinate in 2-D. During the transformation, multiresolution strategies are applied by hierarchical strategies and resizing the image continues. The convolution of I_t and I_s with a Gaussian kernel $G(\sigma_r)$ are defined.

$$I_t(x, r) = G(\sigma_r) \times I_t(x) \tag{43}$$

$$I_s(x, r) = G(\sigma_r) \times I_s(x) \tag{44}$$

Then, different multiresolution strategies are further defined by computing cost function \mathbb{C} in each resolution level ϵ .

$$\mathbb{C}_\epsilon = \sum_{r=1}^N \mathbb{C}(I_t(x, r), T(I_s(x, r))) \tag{45}$$

where N denotes the number of resolution levels, $\gamma \in [1, N]$ denotes each resolution level of transformation. In each resolution, the Gaussian smoothing with down sampling is applied. Smoothing scales $G(\gamma_r)$ of Gaussian kernel are chosen as follows, when $r = 0$, $\sigma_0 = [40, 40]$; when $r = 1$, $\sigma_1 = [20, 20]$; when $r = 2$, $\sigma_2 = [10, 10]$; when $r = 3$, $\sigma_3 = [5, 5]$.

To produce an optimal transformation function, the similarity measure method in this study is calculated by normalized mutual information (MI) [26].

$$MI(\mu; I_t, I_s) = \sum_{x \in L_t} \sum_{y \in L_s} p(t, s; \mu) \log_2 \frac{p(t, s; \mu)}{p(t)ps(s; \mu)} \tag{46}$$

where L_t and L_s are sets of regularly spaced intensity bin centers, p is the discrete joint probability, and pt and ps are the marginal discrete probabilities of the target and source image, obtained by summing p over t and s , respectively. The joint probabilities are estimated with B-spline Parzen windows.

$$p(I_t, I_s; \mu) = \frac{1}{|\tau_t|} \sum_{x_i \in \tau_t} \omega_t \left(\frac{1}{\sigma_f} - \frac{I_t(X_i)}{\sigma_f} \right) \times \omega_s \left(\frac{1}{\sigma_s} - \frac{I_s(X_i)}{\sigma_s} \right) \tag{47}$$

where ω_t and ω_s represent the B-spline Parzen windows of the target and source images. The scaling constants σ_f and σ_s must equal the intensity bin widths defined by L_t and L_s . These follow directly from the gray-value range of I_t and I_s and the user-specified number of histogram bins $|L_t|$ and $|L_s|$. Based on the definition of MI in Eq. 46, the normalized mutual information (NMI) is defined.

$$NMI = \frac{H(I_t) + H(I_s)}{H(I_t, I_s)} \tag{48}$$

where H denotes entropy. With the joint probability defined in Eq. 46, NMI is further defined.

$$NMI(\mu; I_t, I_s) = \frac{\sum_{t \in L_t} pt(t) \log_2 pt(t)}{\sum_{t \in L_t} \sum_{s \in L_s} p(t, s; \mu) \log_2 p(t, s; \mu)} + \frac{\sum_{s \in L_s} ps(s; \mu) \log_2 ps(s; \mu)}{\sum_{t \in L_t} \sum_{s \in L_s} p(t, s; \mu) \log_2 p(t, s; \mu)} \tag{49}$$

To solve the optimization problem in Eq. 41, the optimal transformation parameter vector $\hat{\mu}$ and an iterative optimization strategy is employed.

$$\mu_{k+1} = \mu_k + a_k d_k, k = 0, 1, 2, \dots, \tag{50}$$

where d_k denotes the search index at iteration k , and a_k denotes a scalar gain factor controlling the step size along the search direction. In this study, we apply Adaptive Stochastic Gradient Descent (ASGD) [27] in the elastic transformation. Based on the range of sampler region size (200 pixels \times 200 pixels), the optimizer randomly selects the region for deformation in each iteration. In k rounds of iterations, ASGD gradually optimizes the alignment, where the maximum of k is defined as N_i for transformation ($N_i=500$ for affine; $N_i=50$ for b-spline).

$$\hat{\mu}_k = \mu_{k-1} + \gamma(t_k - 1)g_k - 1, k = 1, 2, 3, \dots, N_i \tag{51}$$

$$t_k = \max(0, t_{k-1} + f(-g_{k-1}^{\hat{f}} g_{k-2}^{\hat{f}})), k = 2, 3, 4, \dots, N_i \tag{52}$$

$$\hat{k} = g_k + \epsilon_k, k = 0, 1, 2, \dots, N_i \tag{53}$$

where $\gamma(t_{k-1})$ denotes the function of the step size at iteration k , g_k is an approximation of the true derivative $g = \partial \mathbb{C} / \partial \mu$ at μ_k , ϵ_k is approximation error and f is a sigmoid function.

For transformation approaches, Affine allows the deformation of rotation, scaling, translation, and shearing.

$$T_\alpha(x) = RGS(x - c) + t + c \tag{54}$$

where c denotes the central point of rotation, t denotes translation vector, and R , G , and S respectively denote the rotation, shearing and scaling matrix.

Compared with affine, b-spline allows the degree of non-rigid deformation through the mesh of control points, which obtains better alignment of the structural change in the transformation to fix the geometric difference of original input datasets.

$$T_b(x) = x + \sum_{x_k \in N_x} p_k \beta^3\left(\frac{x - x_k}{\sigma}\right) \quad (55)$$

where x_k denotes the control points on a regular grid overlaid on the target image, $\beta^3(x)$ denotes the cubic multidimensional B-spline polynomial [28], p_k denotes the B-spline coefficient vectors, σ denotes the B-spline control point spacing and N_x denotes the set of all control points within the compact support of the B-spline at x . In B-spline, the control point grid is defined by the amount of space between the control points $\sigma = (\sigma_1, \dots, \sigma_d)$ (d denotes the dimension of image) and the transformation of a point can be computed from surrounding control points. This is beneficial for modeling local transformations and fast computation.

2.8 Difference Analysis

For difference analysis, the fusion result is generated by overlapping the registered source image to the target image. In order to highlight the difference between two images, the system subtracts and reallocates the pixel intensity to an eight bit range (0-255). The overlapped images I_d is computed from the subtraction of the target image and the registered image. I'_d denotes the fusion result in the intensity range of eight bits.

$$I'_d = 255 \times \frac{I_d - I_{min}}{I_{max} - I_{min}} \quad (56)$$

where $I_{max} = \max(I_d)$ and $I_{min} = \min(I_d)$. The area in the fusion result with high intensity belongs to the target image, and the area with low intensity belongs to the region of the registered source image.

3 Experimental Results

In this study, 142 images from 106 patients were collected from three databases, including the Open-i database (the national library of medicine, USA) [29] for building AI training models, the Chest X-ray8 database (NIH Clinical Center, USA) [30] and a clinical data set collected from the Wan Fang Hospital, Taiwan (IRB-LN201703078) for validating the proposed method. A challenging clinical dataset with the Scoliosis (S-shaped spine) condition as shown in Fig. 9 from the Wan Fang Hospital, Taiwan is selected to test the system's robustness.

To build training models for lungs, ribs and clavicles in the proposed system, 70 frontal-view chest X-ray images of 70 patients, originally from the Indiana University Chest X-ray Collection [31], are selected from the Open-i database. Open-i is an open-access and diverse database, totally containing 7,470 chest X-ray images of 3,851 patients. The selection of the training database is to collect images with varying bone structure appearances, diverse image contrast based on histogram distribution and intensity strength. On the other hand, to validate the performance and implications for patient care of the proposed system, testing images are randomly selected from the datasets collected from the Wan Fang Hospital and an open database, i.e. Chest X-ray8 datasets, which comprises 108,948 frontal-view chest X-ray images of 32,717 patients (from 1992-2015). All images in the Chest X-ray8 are collected from the clinical PACS database at the National Institutes of Health Clinical Center and rescaled to 1024×1024 pixels from the DICOM files. All images from the Wan Fang Hospital are rescaled to 1024×840 - 1248 pixels from the DICOM files. For quantitative evaluation, the average registration error distance is computed among 15 pairs of manually annotated evaluation landmarks. A preliminary test is conducted using 10 pairs of X-ray images (10 patients, 5 male, 5 female; mean age 64 years; range 48-80 years) to compare three elastic registration approaches in the proposed method. Then, a full evaluation is conducted using 36 pairs of chest X-ray images (36 patients, 16 male, 20 female; mean age 49.88 years; range 24-80 years) to compare the proposed method with two current benchmark methods, including BunwarpJ [32] and Fully Automatic Elastic Registration (FAER) [33]. Table 1 shows the data distribution with respect to the data source, the number of patients and images for training and testing. For training, 70 patients' images from the Open-i database (the National Library of Medicine, USA) [29] are collected. For testing, 36 chest X-ray images of 18 patients from the Chest X-ray8 database (NIH Clinical Center, USA) [30] and 36 images of 18 patients from the Wan Fang Hospital, Taiwan (IRB-LN201703078) are utilized. Separate training and testing sets are used for evaluation to guarantee the model is never trained and validated on the same data. The dataset from the Wan Fang Hospital, along with manual annotations, are made available (see the Declaration section at the end of the paper).

3.1 Preliminary Test

In the preliminary test, the quantitative analysis (Table 2) and pair-sample t test (Table 3) are conducted to find the optimal transformation model of the proposed method by comparing three elastic transformation methods. Proposed methods 1, 2 and 3 represent the proposed method with three elastic transformation methods, i.e., affine (Proposed

1), b-spline (Proposed 2) and affine+b-spline (Proposed 3), respectively. Table 2 shows that Proposed 3 achieves the lowest mean registration error distance (MRED) (8.03 mm, 23.38 pixel) compared to Proposed 1 (8.86 mm, 25.79 pixel) and Proposed 2 (10.89 mm, 31.55 pixel), and the lowest mean registration error ratio (MRED) w.r.t. the length of image diagonal (1.46%) compared to Proposed 1 (1.61%) and Proposed 2 (1.98%). Table 2 also shows the computational efficiency of the three proposed methods. For automatic landmark detection and registration of a pair of chest x-ray images, Proposed 1 takes 23.20 seconds; Proposed 2 takes 8.67 seconds; Proposed 3 takes 23.69 seconds. For registration accuracy, Table 3 shows that Proposed 3 obtains significantly better result than Proposed 1 ($P = 0.013$) and Proposed 2 ($P = 0.029$). Thus, Proposed 3 is adopted for further full evaluation.

3.2 Full Evaluation

In full evaluation, the quantitative analysis (Table 4) and paired-sample t test (Table 5) are conducted to compare the proposed method with two benchmark methods, including BunwarpJ and FAER. Table 4 shows that the proposed method achieves the lowest MRED (8.99 mm, 23.54 pixel) compared to BunwarpJ (15.64 mm, 40.97 pixel) and FAER (180.5 mm, 472.69 pixel) and the lowest MRED (1.61%) compared to BunwarpJ (2.81%) and FAER (32.51%), and the computational time of the proposed method (23.69s) is slightly slower than BunwarpJ (12.86s) and FAER (8.25s) but still in seconds. Table 5 shows that the proposed method achieves significantly better results than BunwarpJ ($P = 0.001$) and FAER ($P < 0.001$).

Elastic transformation, including B-spline transformation, has been demonstrated to be effective for soft tissue alignment; however, it may be too flexible in introducing unrealistic deformation at times. The experimental results show that the proposed method, which combines rigid transformation and elastic transformation, is demonstrated to be better than the benchmark elastic registration methods. As in the proposed framework, the hybrid L-SVM and ADAM methods render anatomical landmarks and force constraints for global bone structure (hard tissue) alignment, and afterwards elastic transformation is applied for local registration to refine soft

tissue alignment locally. Figure 7 shows the fusion results of four pairs of chest x-ray images (36 in total) before and after registration generated by Proposed 3 and two existing benchmark approaches.

3.3 Error Distance of 15 Evaluation Landmarks

The mean registration errors of individual evaluation landmarks show that the proposed method achieves the lowest error distance in every evaluation landmark (Fig. 8). The landmark L11 and L7, located near the bottom corner of the lungs, tend to have high registration errors among all methods, as the lung size is retractable depending on the breathing status of the patient during chest radiography. To avoid the influence of data outliers, we removed 14 failed results of FAER, which are given as image width (1024 pixels) in the full evaluation.

3.4 Special Case of Scoliosis

In addition, we evaluated the special case of a scoliosis patient. These images contain high noise and varying structures, making rib and clavicle regions hard to identify. Figure 9 presents this study, which can align the cases of scoliosis, even when the input images are out of the training database.

3.5 System Limitation

Even though the proposed method achieves the most accurate registration result in the evaluation, it rarely fails in few cases, and the proposed method is slower compared to the benchmark method shown in Table 4. Figure 10 presents a failed registration condition due to the interference caused from the medical instrument worn by the patient occluding the target bone features of the clavicle, right lung, and ribs.

4 Discussion and Conclusion

Chest X-rays assist healthcare providers in diagnosing issues that cause symptoms in our heart or lungs, such as difficulty breathing, fever with other signs of infection, pneumonia,

Table 1 Data Distribution of Training and Testing sets

	Database	Location	Images	Patients	Total Patients	Total Images
Training	1. Open-i	National Library of Medicine, USA [29]	70	70	70	70
Testing	2. Chest X-ray8	NIH Clinical Center, USA [30]	36	18	36	72
	3. Hospital	Taipei Municipal Wan Fang Hospital, Taiwan	36	18		
Total					106	142

congestive heart failure, emphysema or chronic obstructive pulmonary disease (COPD), chest pain, chronic cough, lung cancer, and ribcage fracture. During examinations, doctors need to trace the patient's health status for medical diagnosis and evaluation of treatment progress. Tracing a patient's health status is a difficult task, and even experienced doctors can make mistakes, especially when a large number of patients need to be examined. This study presents a fully automatic registration system for chest X-ray images to generate fusion results for difference analysis. Fusion results highlight the difference in the thoracic area, enabling monitoring of patient recovery progress and aiding medical diagnosis and evaluation of treatment progress for thoracic diseases. The proposed system includes data normalization with histogram matching, a hybrid L-SVM model for detection of lungs, ribs and clavicle, an ADMA method to extract corresponding landmarks, a feature-based transformation method (SVD) for coarse global registration and elastic transformation models for local registration. In evaluation, compared with two existing medical image registration methods, the proposed method achieves a significantly lower mean registration error distance ($P \leq 0.001$).

For future work, we would like to investigate the applications of deep learning models in the field of chest X-ray image registration that will decrease the computational time while increasing the accuracy of image registration, and to investigate the applications of the proposed image registration and fusion technology in quantifying and segmenting COVID-19 lung infection and monitoring the treatment

progress of common thoracic diseases, including pneumonia, pneumothorax, atelectasis, cardiomegaly, etc., and further develop quantitative measurements and indicators for clinical applications. In addition, the proposed method may occasionally fail when chest X-ray images are seriously corrupted or image features are heavily occluded by medical instruments. For future technology development, more robust feature detection methods [34] could be integrated into the system to deal with partial lung occlusion.

Acknowledgements This work is supported by the Ministry of Science and Technology of Taiwan (MOST109-2221-E-011 -018 -MY3) and the Wan Fang Hospital funding (107-wf-swf-03).

Data Availability The clinical chest X-rays dataset collected from Wan Fang Hospital, along with the manual annotations, could be downloaded from the GitHub repository (<https://github.com/cwwang1979/Chest-X-ray>), enabling researchers to benchmark their methods and compare them with the proposed method.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

Ethical approval The study was performed following the principles outlined in the Helsinki Declaration and was approved by the Ethics Committee (IRB-LN201703078).

Table 2 The quantitative analysis of error distance of three types of elastic transformation approaches in the proposed method

	MRED (pixel)	MRED (mm)	MRER (%)	N	Std. Dev. (pixel)	Mean Std. Dev. (pixel)	Time (s)
Proposed 1	25.79	8.86	1.61	10	17.32	2.33	23.20
Proposed 2	31.55	10.89	1.98	10	19.55	2.65	8.67
Proposed 3	23.38	8.03	1.46	10	17.62	2.38	23.69

Table 3 The paired-sample t test result comparing the error distance of proposed 3 with proposed 1 and proposed 2

	Paired Differences			95% Mean C. I.		df	P value
	Mean	Std. Dev.	Std. Err. Mean	Lower	Upper		
Prop.3 V.S. Prop.1	-2.41	2.46	0.78	-4.18	-0.65	0.013	23.20
Prop.3 V.S. Prop.2	-8.16	9.95	3.14	-15.29	-1.04	9	0.029

Table 4 The quantitative analysis of error distance comparing the best method in the preliminary test with two benchmark methods

	MRED (pixel)	MRED (mm)	MRER (%)	N	Std. Dev. (pixel)	Mean Std. Dev. (pixel)	Time (s)
Proposed 3	23.54	8.99	1.61	36	22.09	3.68	23.69
Bunwarj [31]	40.97	15.64	2.81	36	29.97	4.99	12.86
FAER [32]	472.69	180.5	32.51	36	500.26	83.37	8.25

Table 5 The paired-sample t test result comparing the error distance of the best method in the preliminary test with two benchmark methods

	Paired Differences			95% Mean C. I.			
	Mean	Std. Dev.	Std. Err. Mean	Lower	Upper	df	P value
Prop.3 V.S. BunwarpJ [31]	-17.42	29.84	4.97	-27.52	-7.32	35	0.001
Prop.3 V.S. FAER [32]	-449.15	495.54	82.59	-616.82	-281.48	35	<0.001

References

- Derek, L. G. H., Philipp, G. B., Mark, H., & David, J. H. (2001). Medical image registration. *Physics in Medicine & Biology*, 46(3), R1.
- Comeau, R. M., Fenster, A., & Peters, T. M. (1998). Intraoperative US in interactive image-guided neurosurgery. *Radio Graphics*, 18(4), 1019–1027.
- Weese, J., Penney, G. P., Desmedt, P., Buzug, T. M., Hill, D. L. G., & Hawkes, D. J. (1997). Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. *IEEE Transactions on Information Technology in Biomedicine*, 1(4), 284–293.
- Khamene, A., Bloch, P., Wein, W., Svatos, M., & Sauer, F. (2006). Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. *Medical Image Analysis*, 10(1), 96–112.
- Hill, D. L., Hawkes, D. J., Gleeson, M. J., Cox, T. C., Strong, A. J., Wong, W. L., et al. (1994). Accurate frameless registration of MR and CT images of the head: applications in planning surgery and radiation therapy. *Radiology*, 191(2), 447–454.
- Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11), 977–1000.
- Dean, C. J., Sykes, J. R., Cooper, R. A., Hatfield, P., Carey, B., Swift, S., et al. (2012). An evaluation of four CT-MRI co-registration techniques for radiotherapy treatment planning of prone rectal cancer patients. *The British Journal of Radiology*, 85(1009), 61–68.
- Seldin, D. W., Esser, P. D., Nichols, A. B., Ratner, S. J., & Alderson, P. O. (1983). Left ventricular volume determined from scintigraphy and digital angiography by a semi-automated geometric method. *Radiology*, 149(3), 809–813.
- Carrillo, A., Duerk, J. L., Lewin, J. S., & Wilson, D. L. (2000). Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment. *IEEE Transactions on Medical Imaging*, 19(3), 175–185.
- Rohr, K., Stiehl, H. S., Sprengel, R., Buzug, T. M., Weese, J., & Kuhn, M. H. (2001). Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on Medical Imaging*, 20(6), 526–534.
- Shi, J., Zhang, W., Tang, M., Caselli, R. J., & Wang, Y. (2017). Conformal invariants for multiply connected surfaces: Application to landmark curve-based brain morphometry analysis. *Medical Image Analysis*, 35, 517–529.
- Kalavagunta, C., Zhou, X., Stephen, C. Schmechel., & Gregory, J. Metzger. (2015). Registration of in vivo prostate MRI and pseudo-whole mount histology using Local Affine Transformations guided by Internal Structures (LATIS). *Journal of Magnetic Resonance Imaging*, 41(4), 1104–1114.
- Rivest-Hénault, D., Dowson, N., Greer, P. B., Fripp, J., & Dowling, J. A. (2015). Robust inverse-consistent affine CT-MR registration in MRI-assisted and MRI-alone prostate radiation therapy. *Medical Image Analysis*, 23(1), 56–69.
- Gu, S., Meng, X., Scieurba, F. C., Ma, H., Leader, J., Kaminski, N., et al. (2014). Bidirectional elastic image registration using B-spline affine transformation. *Computerized Medical Imaging and Graphics*, 38(4), 306–314.
- Szeliski, R., & Coughlan, J. (1997). Spline-Based Image Registration. *International Journal of Computer Vision*, 22(3), 199–218.
- Thirion, J. P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3), 243–260.
- Zhang, K. Y. J., & Main, P. (1990). Histogram matching as a new density modification technique for phase refinement and extension of protein molecules. *Acta Crystallographica Section A*, 46(1), 41–46.

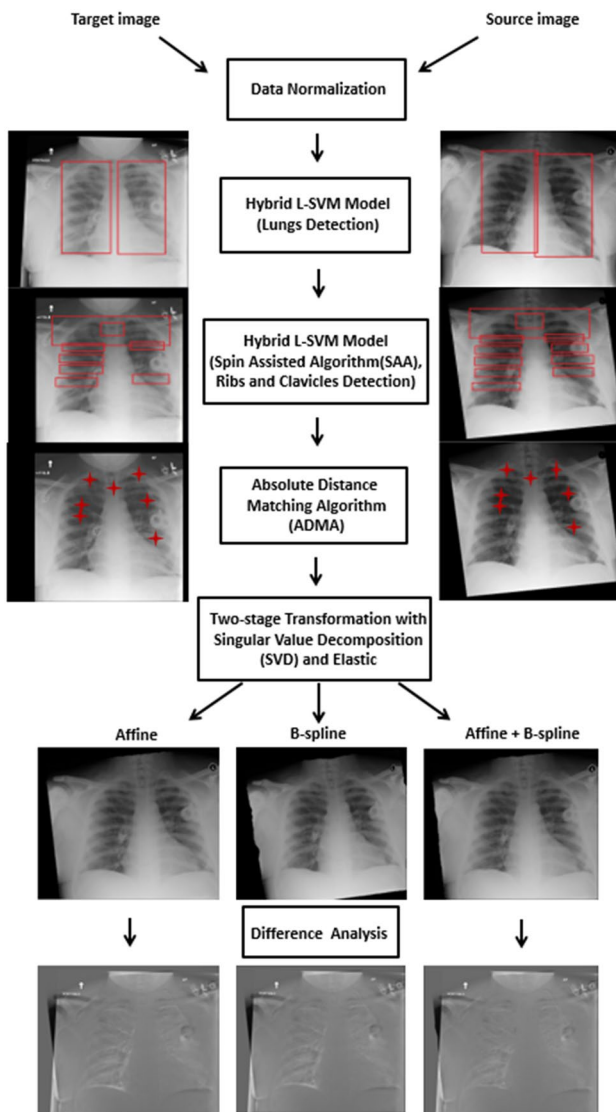


Fig. 1 The flow chart of the proposed method

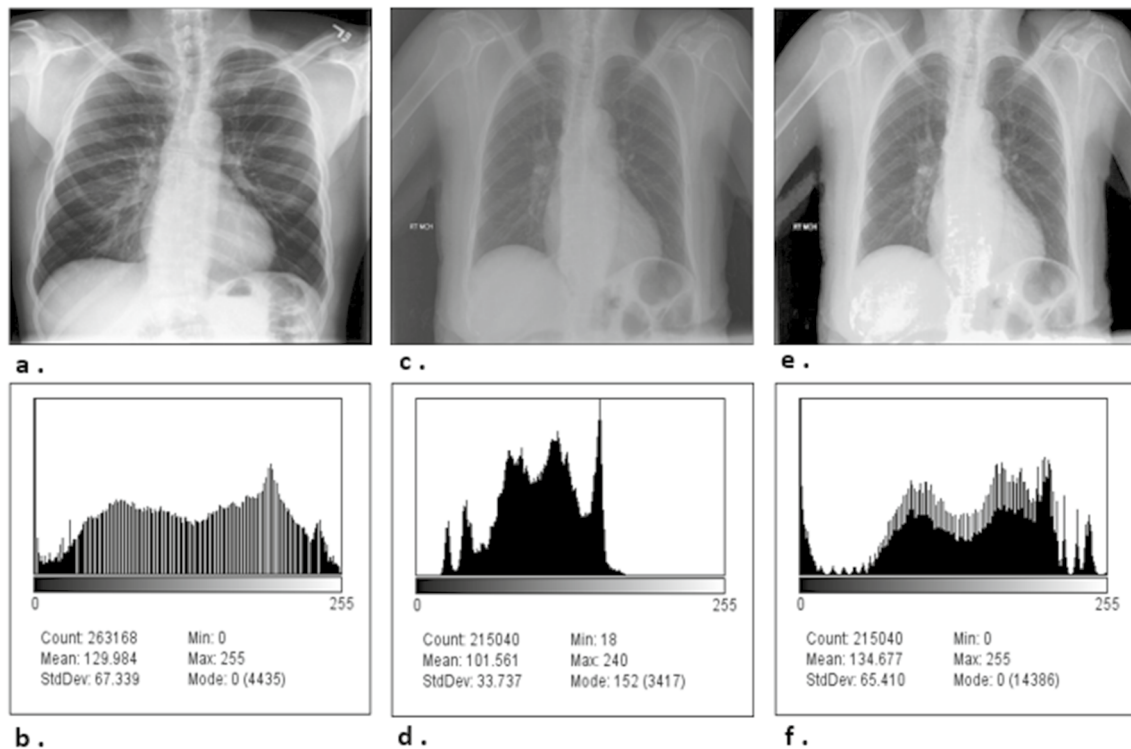


Fig. 2 Illustration of histogram matching. **a** The reference image. **b** Histogram of reference image. **c** One of the input images for histogram matching. **d** Histogram of input image. **e** Input image after histogram matching. **f** Histogram of input image after histogram matching

18. Helmer, E. H., & Rufenacht, B. (2005). Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching. *Photogrammetric Engineering Remote Sensing*, 71(9), 1079–1089.
19. Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9), 1627–1645.
20. Li, B., Cheng, K., & Yu, Z. (2016). Histogram of Oriented Gradient Based Gist Feature for Building Recognition. *Intell. Neuroscience*, 2016, 10.
21. Dalal, N., & Triggs, B. “Histograms of oriented gradients for human detection.” pp. 886–893 vol. 1.
22. Arun, K. S., Huang, T. S., & Blostein, S. D. (1987). Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5), 698–700.
23. Shinji, U. (1991). Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13, 376–380.
24. Klein, S., Staring, M., Murphy, K., Viergever, M. A., & Pluim, J. P. W. (2010). Elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*, 29(1), 196–205.
25. Shamonin, D., Bron, E., Lelieveldt, B., Smits, M., Klein, S., & Staring, M. (2014). “Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer’s Disease,” *Frontiers in Neuroinformatics*, vol. 7, no. 50, 2014-January-16.
26. Thevenaz, P., & Unser, M. A. (1997). “Spline pyramids for inter-modal image registration using mutual information” *Proc. SPIE 3169, Wavelet Applications in Signal and Image Processing V*, 30 October.
27. Klein, S., Pluim, J. P. W., Staring, M., & Viergever, M. A. (2008). Adaptive Stochastic Gradient Descent Optimisation for Image Registration. *International Journal of Computer Vision*, 81(3), 227.
28. Unser, M. (1999). Splines: a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6), 22–38.
29. “Openi - an open access biomedical image search engine. <http://openi.nlm.nih.gov>. Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine.”
30. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. (2017). “ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3462–3471.
31. Demner-Fushman, D., Kohli, M. D., Marabini, R., Rosenman, M. B., Shooshan, S. E., Rodriguez, L., Thoma, G. R., & McDonald, C. J. (2016). Preparing a collection of radiology examinations for distribution and retrieval. *Journal of the American Medical Informatics Association*, 23, 304–310.
32. Arganda-Carreras, I., Sorzano, C. O. S., Marabini, R., Carazo, J.-M., Solorzano, C. O.-d., & Kybic, J. (2006). “Consistent and Elastic Registration of Histological Sections using Vector-Spline Regularization,” *CVAMIA: Computer Vision Approaches to Medical Image Analysis*, vol. 4241, pp. 85–95.
33. Wang, C.-W., Ka, S.-M., & Chen, A. (2014). Robust image registration of biological microscopic images. *Scientific Reports*, 4, 6050.
34. Ouyang, W., & Wang, X. “A discriminative deep model for pedestrian detection with occlusion handling.” pp. 3258–3265.

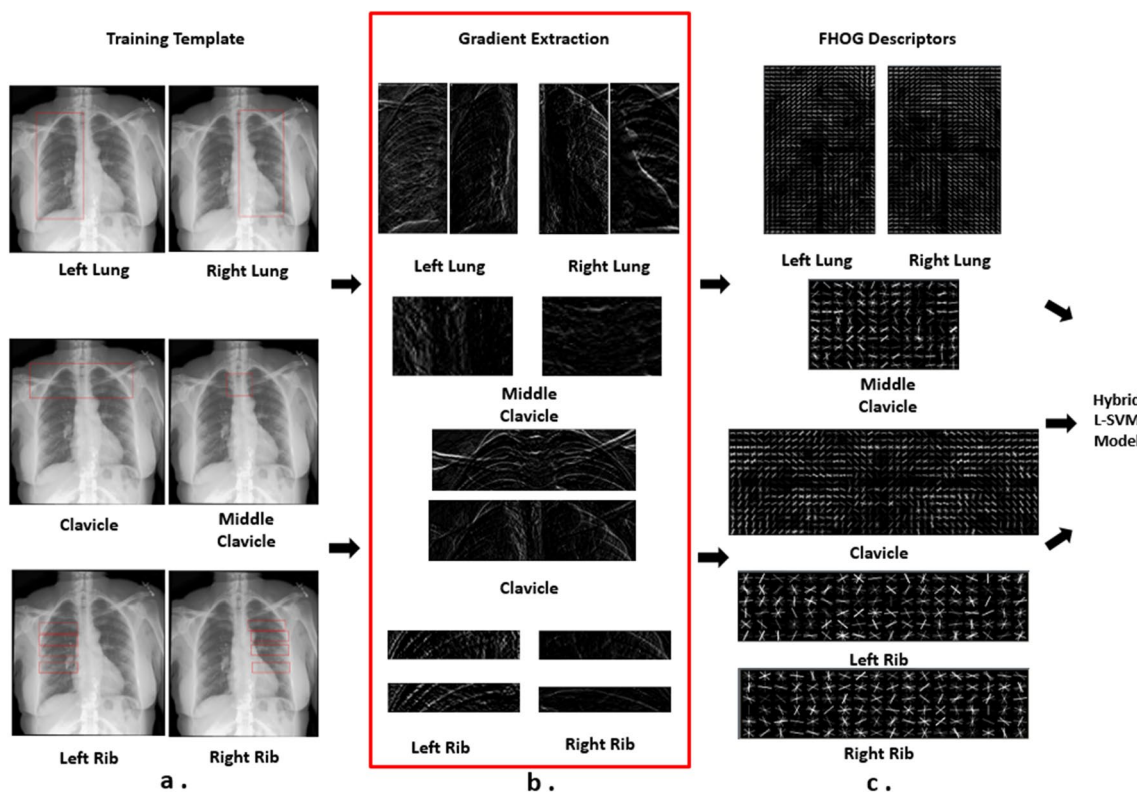


Fig. 3 The illustration from training templates to the hybrid L-SVM model. **a** Training templates manually labeled for lungs, ribs and clavicle. **b** The gradient magnitude of training template is computed

with horizontal and vertical approximations. **c** FHOG descriptor for each L-SVM model is formed by massive 31-dimensional features for classification

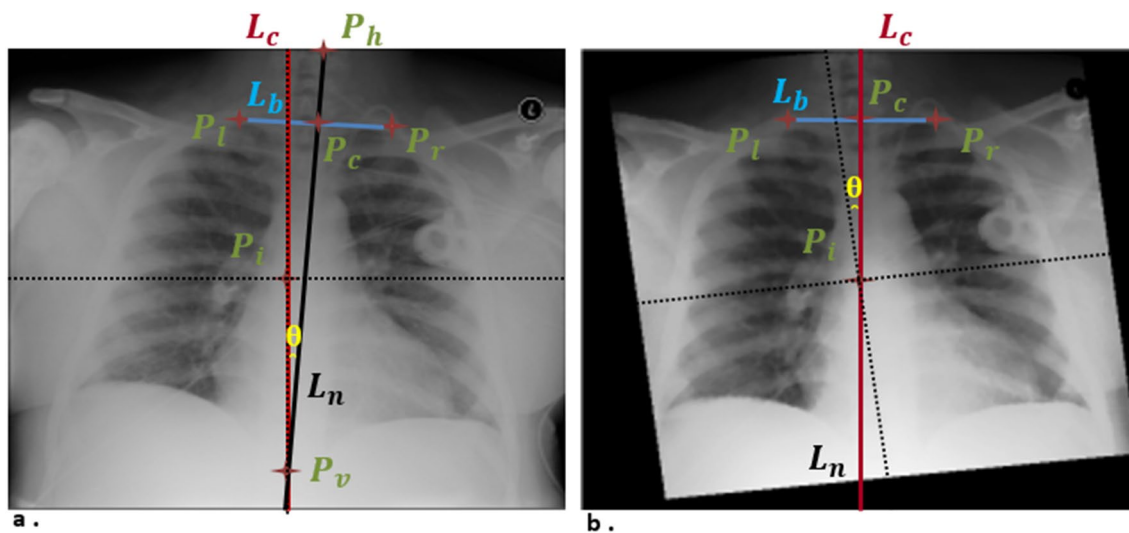


Fig. 4 The illustration to present the definition of rotation angle θ of SAA. **a** The image after lung detection. **b** The modified image after SAA

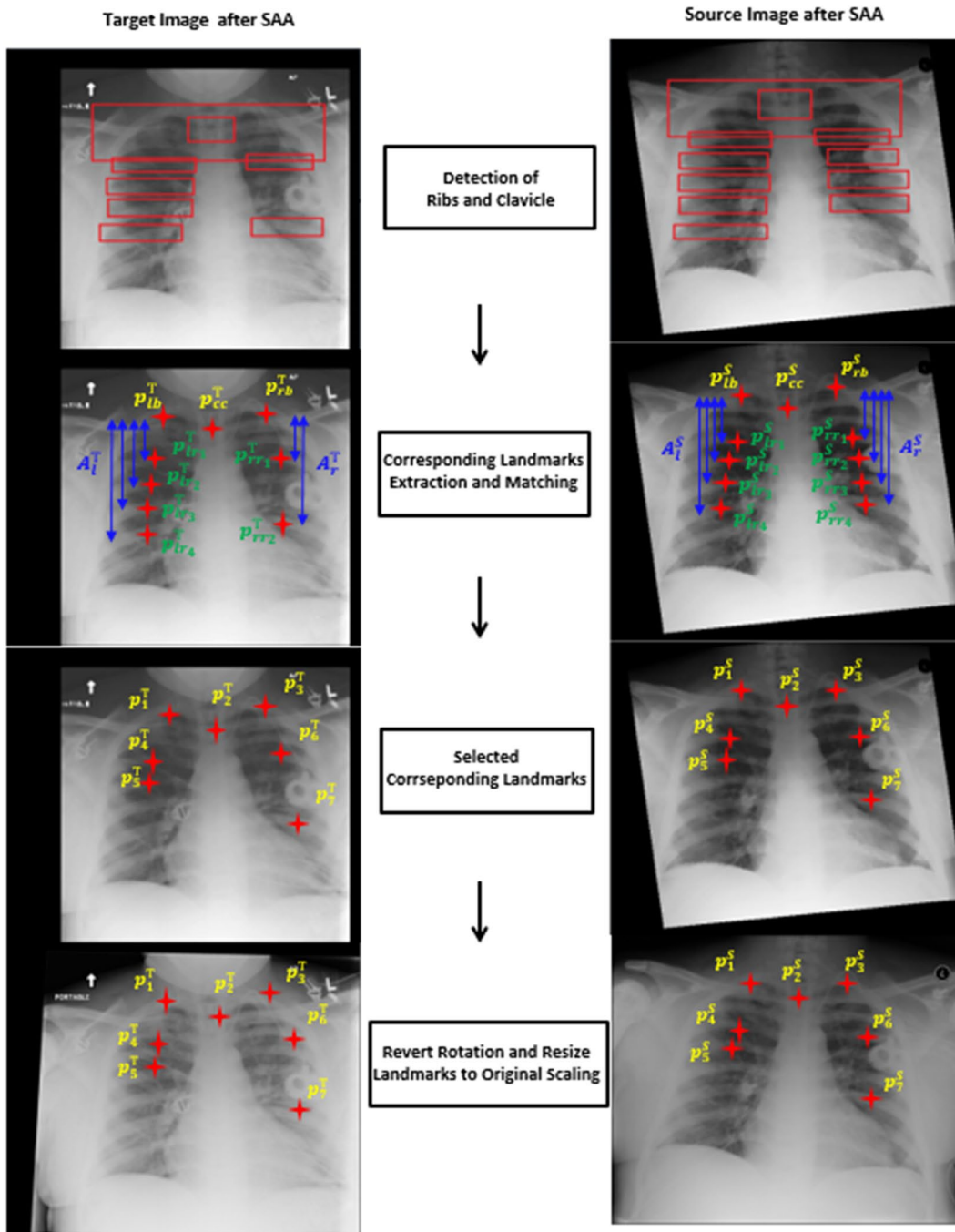


Fig. 5 The illustration of ADMA process

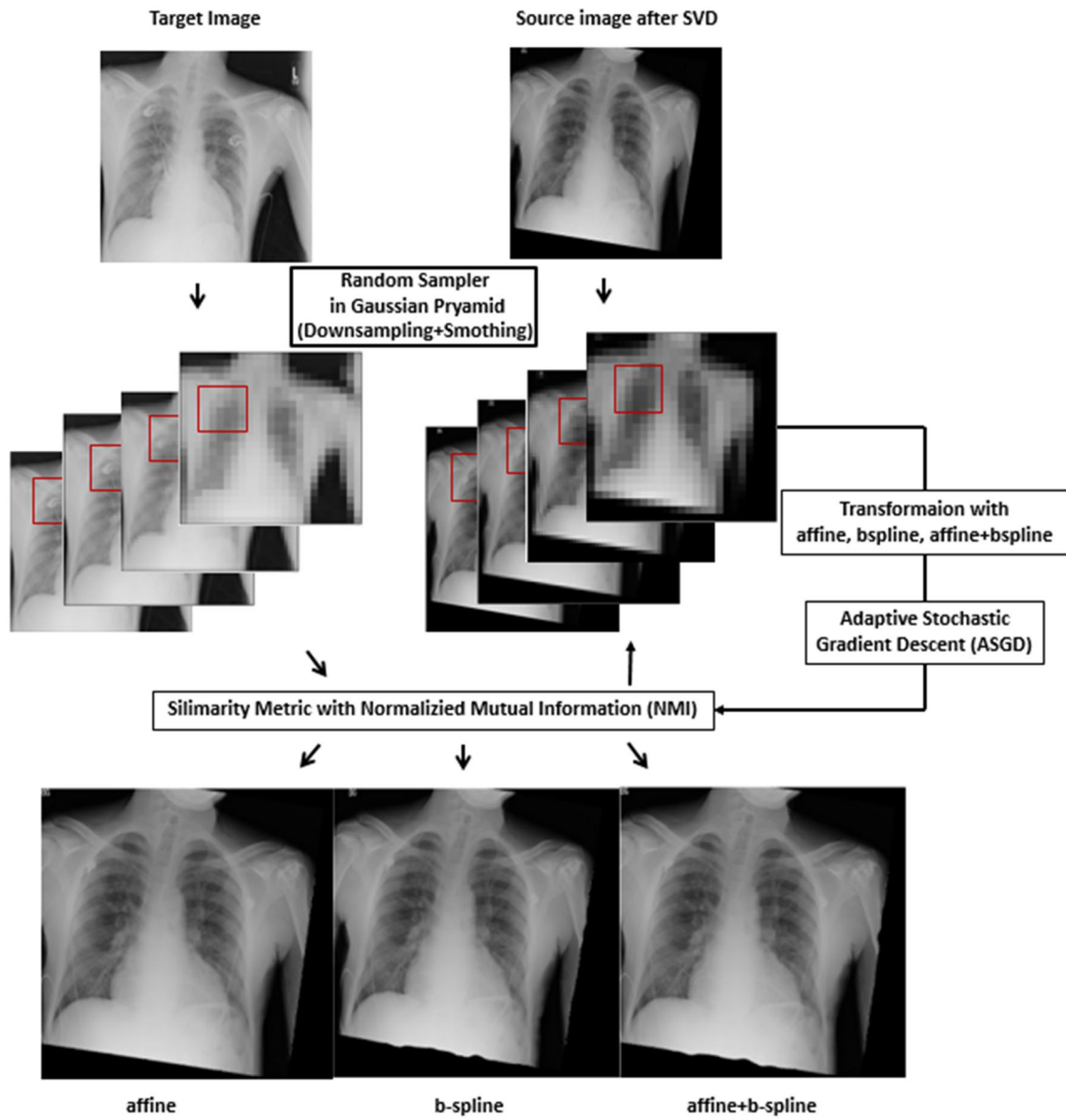


Fig. 6 The illustration to show the process of elastic transformation

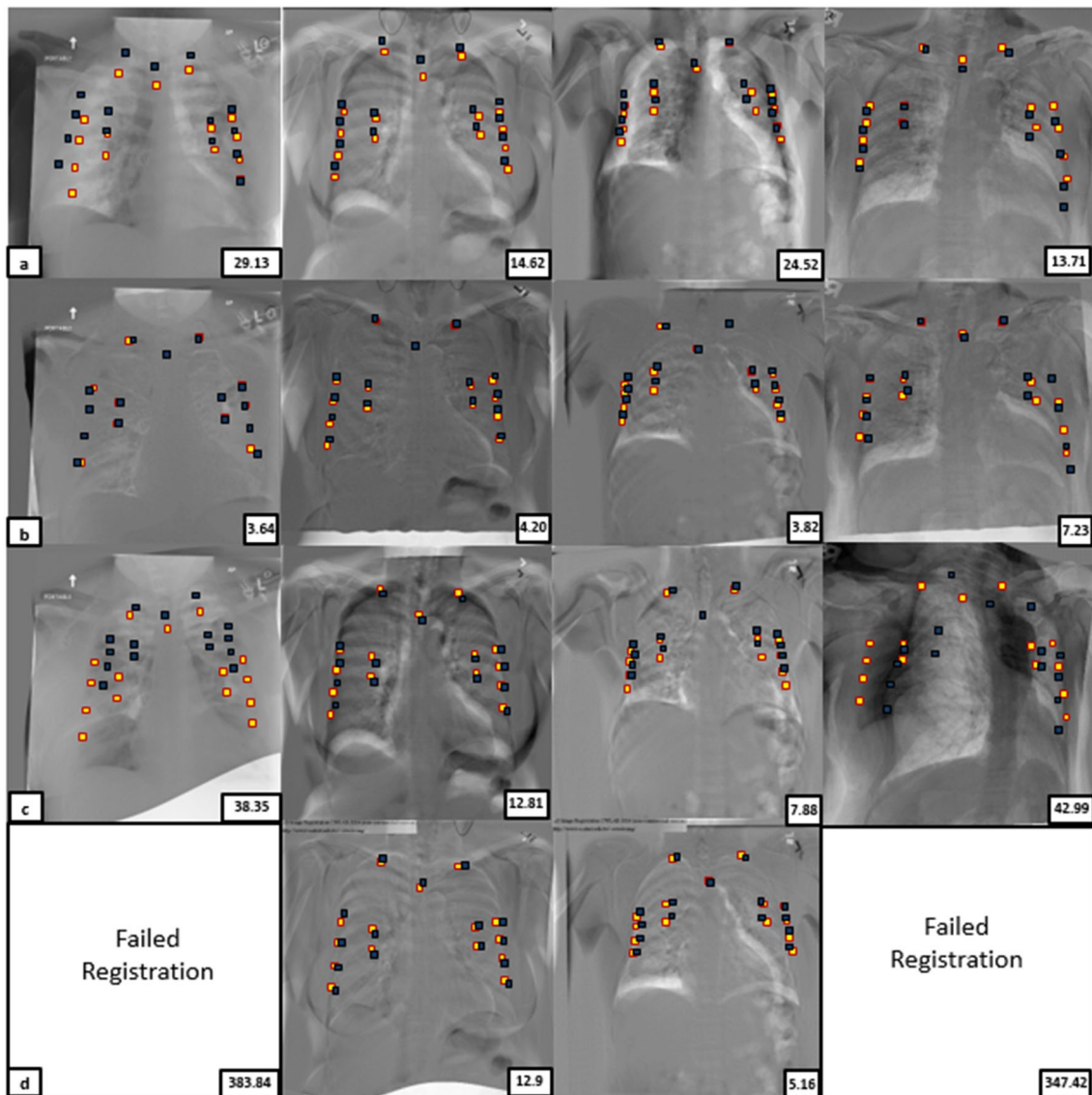


Fig. 7 The difference analysis of four types of testing datasets generated by three methods in the full evaluation. Yellow and blue rectangles separately indicate the location of 15 evaluation landmarks on target and registered source images. The average error distance in millimeter is labeled at the lower right of every difference analy-

sis. Row **a** indicates the difference analysis before registration. Row **b** indicates the difference analysis for the proposed method. Row **c** indicates the difference analysis for BunwarpJ. Row **d** indicates the difference analysis for FAER

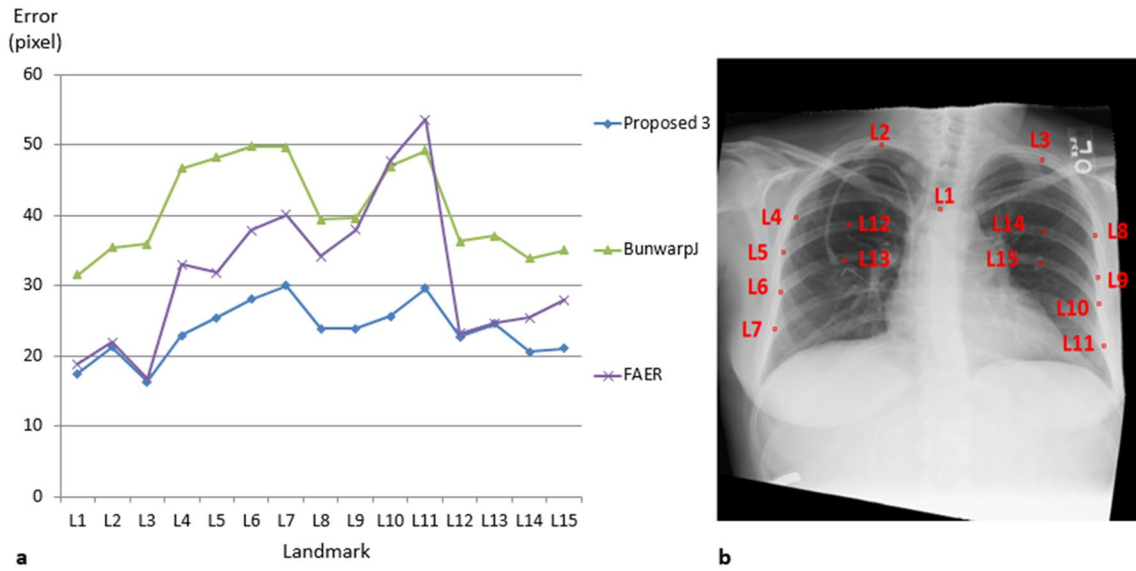


Fig. 8 **a** The mean registration errors of 15 evaluation landmarks for Proposed 3 and two benchmark approaches, showing that Proposed 3 consistently obtains the lowest MRED overall. **b** The layout of 15 evaluation landmarks in a chest X-ray image

Fig. 9 **a** The input datasets of patient suffered from scoliosis. **b** The registered result and the difference analysis generated from the proposed method

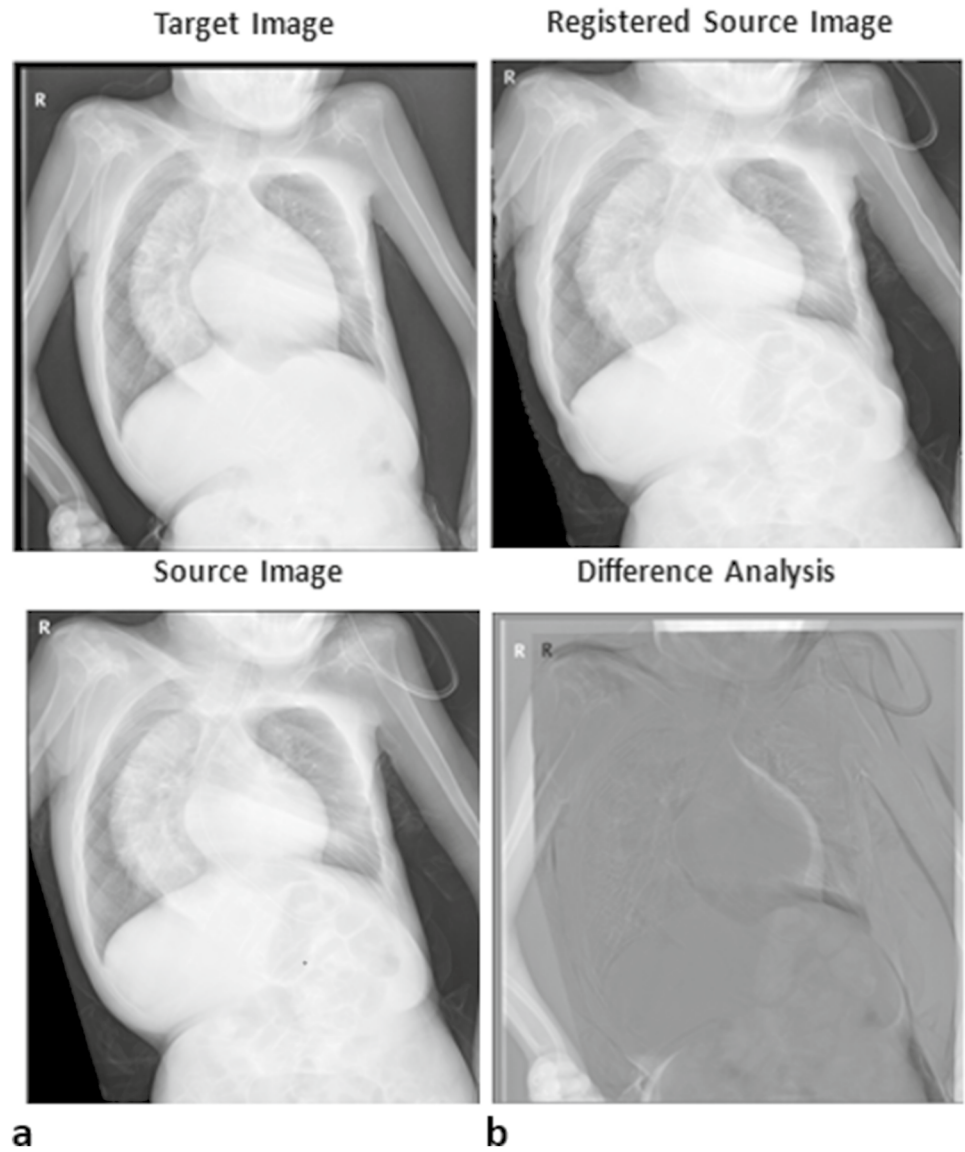


Fig. 10 **a** The input datasets with medical instruments worn by the patient. **b** The failed alignment and the difference analysis generated by the proposed method due to the occlusion of local feature

