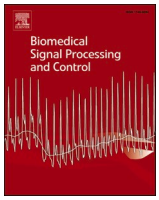
Biomedical Signal Processing and Control 87 (2024) 105505

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Deep learning to predict respiratory lung-tissue displacement from multi-phase pulmonary computed tomography images

Chengmin Zhou [a,b,1], Liyilei Su [a,b,1], Tiegong Wang [c,1], Jinxing Gao [d], Jing Li [e], Jun Liu [d], Zongzhou Ye [d], Xiaoming Zhai [d], Weiguo Xie [d], Chengwei Shao [c,*], Bingding Huang [a,b,*]

[a] College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518188, China
[b] Weiding Joint Laboratory of Medical Artificial Intelligence, Shenzhen Technology University, Shenzhen 518188, China
[c] Department of Radiology, Changhai Hospital, Naval Medical University (Second Military Medical University), Shanghai 200433, China
[d] Wuerzburg Dynamics Inc., Shenzhen 518118, China
[e] Department of Pulmonary and Critical Care Medicine, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, Guangdong, China

ARTICLE INFO

Keywords:
Deep learning
Lung-tissue displacement prediction
Unsupervised learning
Lung registration

ABSTRACT

Lung surgery, particularly lung puncture, is challenging because respiration motion significantly displaces lung tissues (e.g., blood vessels, trachea), including targeted lesions. To improve the success of lung surgery, it is essential to determine how lung tissues move during respiration. Early work used simple mathematical models (e.g., linear models) to predict lung-tissue displacement. However, the quality and quantity of pulmonary computed tomography (CT) images used to build such models were limited, preventing their universal application. Moreover, such mathematical models are too simple to describe how lung tissues are displaced during respiration. In this work, we analyzed the relationship between body-surface and lung-tissue displacement in a large number of multi-phase pulmonary CT images to create a model predicting lung-tissue displacement. First, a systematic method based on unsupervised learning approaches was designed to calculate the displacement of the body surface and lung tissues. Second, a deep neural network model was developed to predict lung-tissue displacement, using the body-surface displacement observed in CT imaging data as input. Multi-phase pulmonary CT images from 199 patients were collected to train and test our model. The validation experiment demonstrated that the prediction accuracy based on our deep learning model reached 79.8%, which outperformed traditional machine learning approaches like the linear model (79.2%), linear support vector regression model (79.2%), and support vector regression model with radial basis function (67.2%). These results demonstrated that our method could predict lung-tissue displacement based on body-surface displacement during respiration, potentially improving the success of lung puncture procedures.

1. Introduction

Lung cancer is one of the most frequently occurring and deadly cancers in the United States, making it an important public health concern [1]. Diagnosing lung lesions as benign or malignant influences the treatment strategy and prognosis of the disease [2,3]. Biopsy and ablation of pulmonary lesions by percutaneous puncture are highly accurate for diagnosing malignant chest diseases and play a crucial role in lung cancer treatment [4]. Precise navigation to the lesion is essential for pulmonary puncture biopsy and local treatment. However, in practice, a patient's respiratory movements can cause significant deviation of the lesion, requiring multiple adjustments to the direction and depth of the puncture needle to reach the target site [5]. Clinical studies show that the lungs, liver, and kidneys are most affected by pulmonary respiratory movements [6]. Therefore, there is an urgent need to study the effect of respiratory motion on lesion displacement to design more accurate localization and puncture guidance techniques and improve lung surgery's success rate.

Many studies have been conducted to reduce the impact of respiratory movement during lung surgery [7,8]. In the field of puncture procedure, Zheng et al. [9] proposed a novel follow-up compensatory method for respiratory movement in robot-assisted puncture

* Corresponding authors at: College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518188, China.
E-mail addresses: cwshao@sina.com (C. Shao), huangbingding@sztu.edu.cn (B. Huang).
[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.bspc.2023.105505
Received 25 May 2023; Received in revised form 14 August 2023; Accepted 23 September 2023
Available online 30 September 2023
1746-8094/

procedures. Ma et al. [10] proposed a framework based on long short-term memory (LSTM) to correlate the movement of body-surface markers with the displacement of internal tumor lesions to more accurately predict real-time tumor position during robot-assisted biopsy procedures.

Both direct and indirect measurements can measure lung-tissue displacement in real time. Direct measurement is more accurate, but it may damage the lung tissues. Indirect measurement (e.g., estimation based on body-surface movement) may introduce error, although it avoids damaging the tissues. Indirect modeling can be implemented by different approaches such as linear regression [11], non-linear polynomial regression [12], B-spline function [13], principal component analysis [14], fuzzy logic [15], support vector regression (SVR) [16], and deep neural networks [17]. Due to the disadvantage of taking direct measurements during surgery, there are few lung computed tomography (CT) images to build models from, which is insufficient for universal application. Therefore, indirect measurement would be preferred if the measurement error can be eliminated or constrained to an acceptable range. However, existing mathematical models are simple and thus cannot precisely describe how lung tissues are displaced during respiration movement.

Point-set registration with optical markers is widely used with CT imaging to obtain lung-tissue displacement measurements at different respiration phases, allowing a relationship to be determined between marker displacement on the CT images and the actual displacement of the patient's lung. In other words, image coordinates are mapped to the physical coordinates [18]. Some image-to-patient registration methods have been proposed [19,20], which include custom optical markers automatically positioned in the patient's physical space. However, as the positions of these markers affect the registration result, finding optimized marker placement positions that are universally applicable and safe for patients is a secondary aim of this study.

In this paper, we analyse the relationship between body-surface displacement and lung-tissue displacement. First, we propose a method to model and track respiratory lung-tissue movement using optical body-surface markers. This method starts with designing an initial pulmonary CT data collection plan. The initial collection plan investigates CT images taken from two phases of the respiratory cycle (the end of inspiration and the small inspiration phase). These data are used to generate the first set of displacement values of lung-tissue and body surface, which are used for unsupervised learning approaches (clustering). The clustering results are then used to generate optimized optical marker placement positions and optimize the data collection plan. The optimized collection plan is used to collect new CT imaging data taken from three phases of the respiratory cycle (the end of expiration, the small inspiration phase, and the end of inspiration). A new set of displacement values is generated from this data, which is used for regression (i.e., training the model). Finally, a neural network is designed and trained to predict lung-tissue displacement utilizing several multi-phase pulmonary CT images. Moreover, three traditional machine-learning approaches are also tested for comparisons.

2. Methods

In this work, we used indirect measurements to predict lung-tissue displacement from different phases of pulmonary CT images with optical markers. To reduce prediction error, we employed: (1) a large CT imaging dataset, (2) a systematic method based on unsupervised learning, and (3) a deep neural network.

Fig. 1 shows a schematic representation of our proposed method, which follows these steps:

(1) Lung data 1 and surface data 1 of two respiration phases are collected based on the initial data collection plan.

(2) Lung data 1 and surface data 1 are used to generate Lung-tissue displacement 1 and surface displacement 1 respectively after data processing.

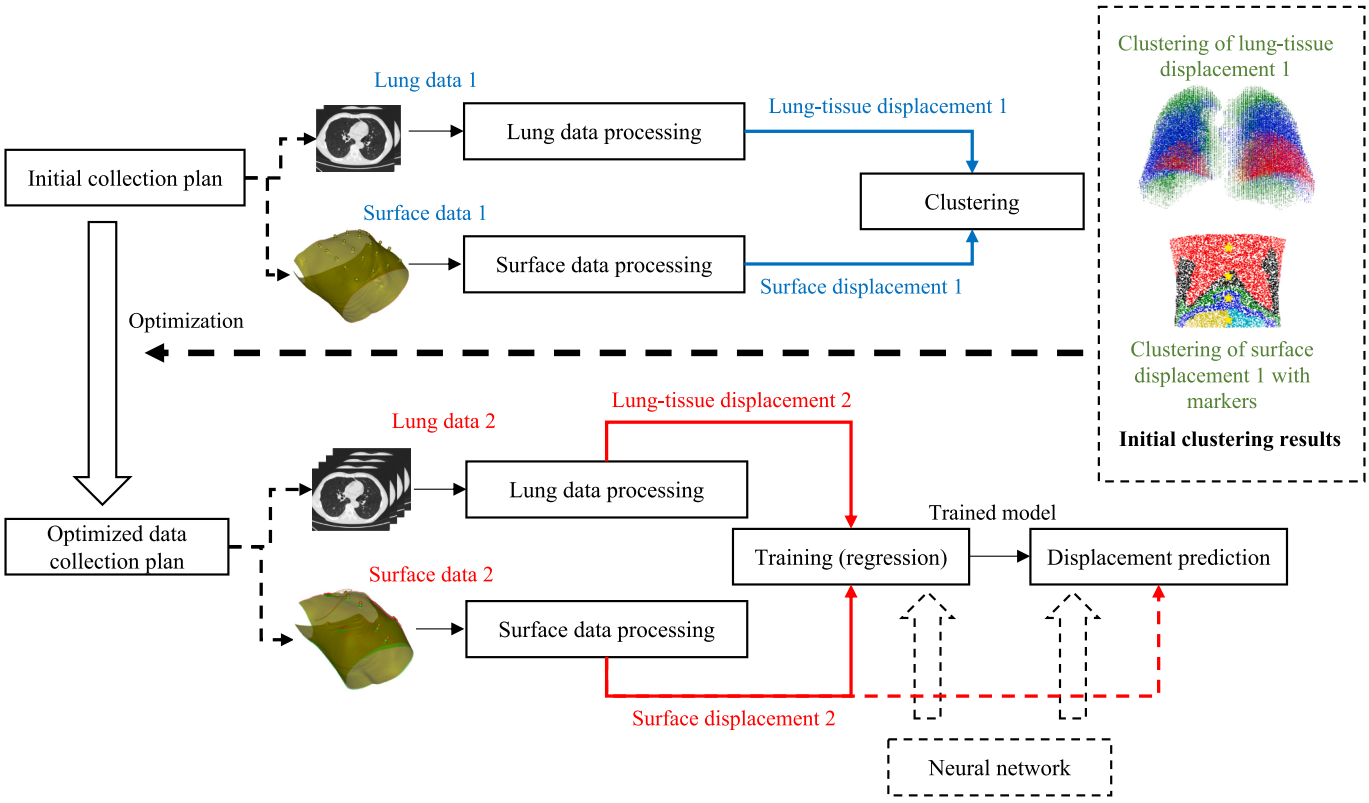


Fig. 1. The workflow of the proposed displacement prediction method. First, the initial collection plan is used to acquire the initial clustering results. The initial clustering results contribute to the optimization of the initial data collection plan, resulting in *an optimized data collection plan.*

(3) The lung and surface data are sorted into different classes by an unsupervised learning approach (clustering), according to the intensity of movement obtained from the registration. Based on the clustering results, optimized optical marker placement positions are determined.

(4) The clustering results contribute to optimizing the data collection plan, which is used to collect new data (lung data 2 and surface data 2) from three respiration phases.

(5) New displacements (lung-tissue displacement 2 and surface displacement 2) are obtained.

(6) A neural network is designed and trained for displacement prediction. The training inputs are lung-tissue displacement 2 and surface displacement 2, while the prediction input is surface displacement 2.

2.1. Initial data collection plan

The initial data collection plan aims to determine the movement intensities of the lung and body surface, thereby determining the marker placement and optimizing the data collection plan. Pulmonary CT images from two respiration phases were used to calculate the displacements. Initially, these phases were at the end of the inspiration and the small inspiration phase, respectively. These phases were chosen for two reasons: (1) it is easier for patients to hold their breath at the end of inspiration, and (2) larger displacement is obtained with these phases, as shown in Fig. 2.

2.2. Data processing

Data processing included lung segmentation, 3D lung reconstruction, down-sampling of lung data and reconstruction, and down-sampling of surface data. The data processing was performed before computing the lung and surface displacements.

2.2.1. Lung data processing

Lung segmentation. We applied lung segmentation to remove noise (i.e., other body tissues) from initial lung CT images. The segmentation procedure consisted of four steps, shown in Fig. 3: (1) preprocessing of CT image, (2) initial lung segmentation, (3) trachea removal, and (4) division of left and right lungs.

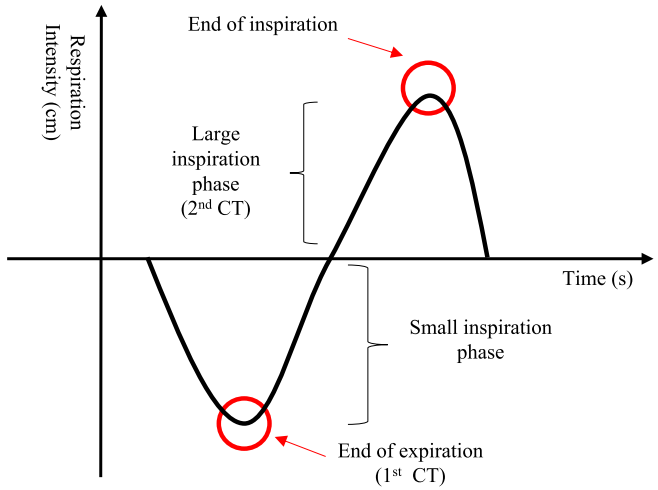


Fig. 2. Definitions of the inspiration phases. We define inspiration to be the small inspiration phase and large inspiration phase. The inspiration starts from the end of expiration where the first CT is acquired. When the volunteer or patient relaxes and does not use the muscle to inspirate or respirate, we claim that the lung is at a "balanced state" which is the division of the small inspiration phase and large inspiration phase. The second CT is acquired at the large inspiration phase. We chose the end of expiration and large inspiration phase to collect two CT images because it is easy for the volunteers to hold the breath or control the muscle related to breath. This provides stability and reliability when collecting the CT images.

First, CT image preprocessing denoted noise reduction and image binarization using a curvature flow filter and adaptive thresholding. Then, initial lung segmentation was achieved with a threshold segmentation algorithm combined with morphological algorithms. Next, the trachea was removed using the region-growing method to obtain a trachea mask. Finally, the projection-integration method and the morphological algorithm divided the left and right lungs.

3D lung reconstruction. After the lung segmentation, the marching-cubes algorithm was used to reconstruct the lung 3D images (Fig. 4b-c) from initial pulmonary CT images (Fig. 4a).

Calculation of displacement 1 (lung registration). Multi-resolution B-spline transformation was applied to align two lung CT images using 3D reconstruction of lung contours as regions of interest, adaptive stochastic gradient descent as an optimizer, and linear and B-spline interpolation as interpolation algorithms. Furthermore, the Gaussian pyramid was used to smooth and down-sample the images in multi-resolution transformation. After registration, the displacement field of the lung was extracted. (Fig. 4d).

Down-sampling. The displacement field of the whole lung is in high resolution with millions of vectors, which take too much computing resource in the following processes. Therefore, we randomly sampled 50,000 vectors for the lung, as shown in Fig. 4e.

2.2.2. Surface data processing

Initial CT images were reconstructed to obtain the 3D surface image by selecting the edge pixels (Fig. 5a). The processing of Surface data 1 shares some steps with that of Lung data 1: 3D reconstruction (Fig. 5a), registration (Fig. 5b), and Surface displacements 1 represented by vectors (Fig. 5c). After these steps, the surface image after down-sampling of 10,000 vectors is obtained (Fig. 5d).

2.3. Clustering and marker placement

The lung is not displaced uniformly and displacement of the whole lung cannot be predicted directly. Therefore, pulmonary CT images were divided into areas using clustering. Each data point in the CT image contains values of both the magnitude and direction of displacement. Grouping data points with similar values together allows the displacement of each area to be predicted.

A small cluster number may result in each resulting area containing a large range of displacement values and therefore cause a significant prediction error. Conversely, a large cluster number is expected to bring inaccuracy because of excessive differences of lungs from different volunteers, resulting in different clustering results in the lungs from different volunteers. This blocks further investigations to find the shared trend of lung movements among volunteers. Lung-tissue displacement was thus clustered by the fuzzy c-means algorithm with the number of cluster centers set to three for the left lung or right lung, dividing the left or right lung into three parts according to the magnitude of displacement (green, blue, red in Fig. 6ab).

Similarly, the fuzzy c-means clustering algorithm was used on the surface-displacement data to determine the placement of the optical markers, with the number of clustering centers set to six, dividing the body-surface displacement into six parts, as shown in Fig. 7a. As the ribs constrain the respiration process, areas where the body surface connects to a rib show a small displacement, as shown in Fig. 7a (red). Other areas of the body surface had more substantial displacement and were organized into five classes, with displacement increasing gradually from the chest to the belly. The surface near the belly formed two classes, represented by light blue and yellow in Fig. 7a. Their displacements were almost the same magnitude but differed significantly in direction.

Within the clustering results, the center of each class (the clustering center) was expected to be an ideal point to place an optical marker to track the respiration motion, as the displacement of the clustering center is usually the average value of this class and is the ideal representation

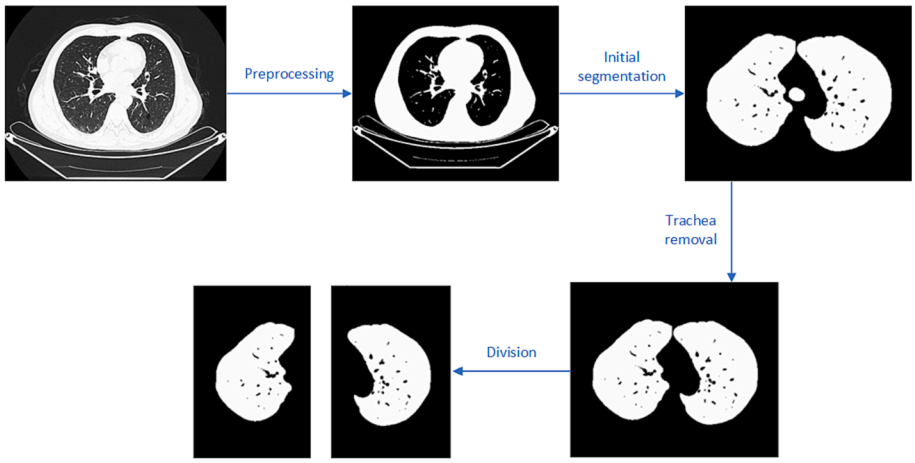


Fig. 3. The mechanism of lung segmentation.

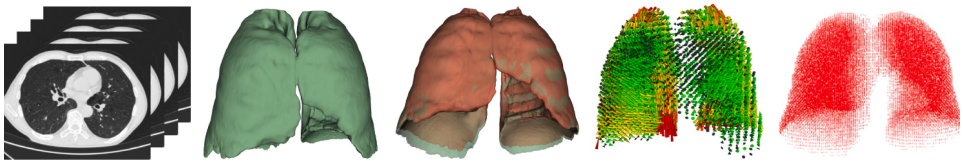


Fig. 4. Lung data processing with the initial data collection plan. (a) Initial pulmonary CT images. (b) Lung 3D image after lung segmentation and reconstruction. (c) Overlapping of two lung 3D images from two phases of respiration, the end of inspiration and the small inspiration phase. (d) Lung-tissue displacements 1 represented by vectors. (e) Lung 3D image after down-sampling.

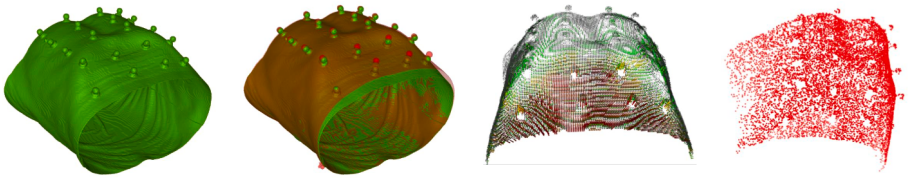


Fig. 5. Results of surface data processing with the initial data collection plan. (a) 3D reconstruction of CT images of small inspiration phase. (b) Overlapping of two surface images from two phases of respiration (the end of inspiration and small inspiration phase). (c) Surface displacements 1 represented by vectors. (d) Surface image after down-sampling.

and feature (input) of regression algorithms. Nevertheless, not all possible positions were suitable. For example, some clustering centers were easy to compute but challenging for humans to distinguish them. The marker placement positions should be distinguishable, stable, and universal, ensuring that data collected from the body surface is reliable. Considering both anatomic knowledge (Fig. 9a) and the clustering results, the final marker placement positions were determined, indicated by yellow stars in Fig. 7b. Four markers were placed on the body surface: one in the center of the chest, one in the xiphoid, and the remaining two markers at points equally spaced between the xiphoid and belly button. These marker positions showed little variation for most patients.

2.4. Optimization of the data collection plan

After analyzing the clustering results with initial data, the optimized data collection plan was confirmed. First, pulmonary CT images were collected from three phases of respiration (the end of inspiration, the small inspiration phase, and the end of expiration), and two lung-tissue displacements were obtained. At the same time, the markers were used to compute the surface data. Comparisons of the initial and optimized data collection plans are presented in Table 1.

2.5. New displacements calculation

The processing of lung data 2 was similar to that of lung data 1. With the CT image data from three respiration phases (end of inspiration, small inspiration phase, and end of expiration), two lung-tissue displacements were computed (end of inspiration-small inspiration phase, end of inspiration-end of expiration) as shown in Fig. 8a and b and the displacement vectors shown in Fig. 8c and d.

Considering the registration result of end of inspiration (green)-end of expiration (red) as an example (Fig. 8e), we noticed that the vessels almost overlapped after registration. This means the registration can simulate the respiration process of humans with less error, assuming the motion of lung tissue is linear. In reality, the simulated trajectory of lung tissue is slightly different from the real trajectory of lung tissue in the respiration process, but we tolerate these differences in lung tissue trajectories because the motion of lung tissue is near linear. These registration results ground the following lung-tissue displacement prediction.

Considering the clustering result of lung-tissue displacement 2

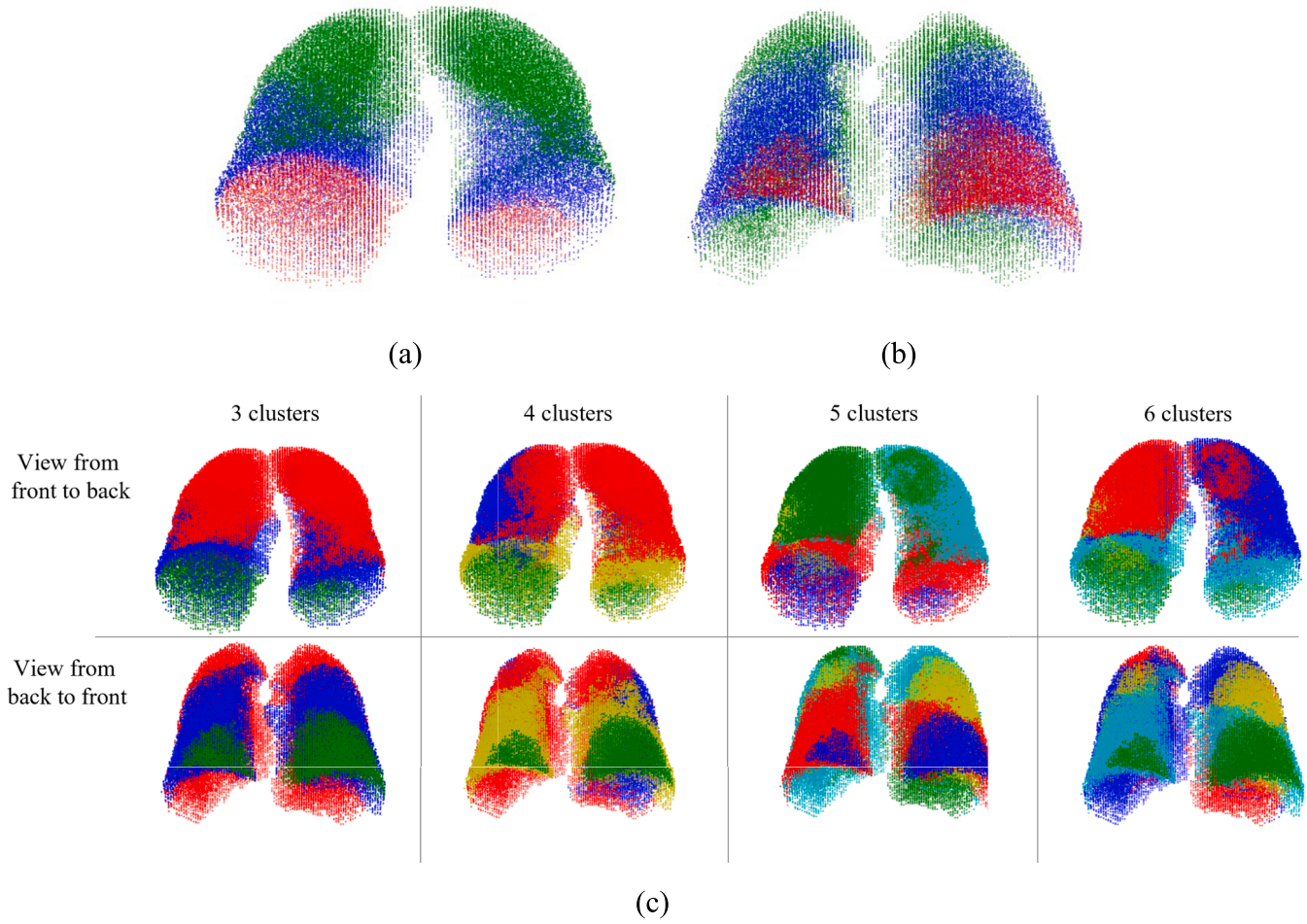


Fig. 6. Clustering results based on Lung-tissue displacements 1 from two viewpoints: (a) front to back and (b) back to front. (c) presents different lung clustering results with 3, 4, 5, and 6 cluster centers respectively. With the increase of clustering centers, more confusing clustering results are acquired. This is expected to block further investigations to find the shared trend of lung movements among volunteers.

(Fig. 6a) and anatomic knowledge of the lung (Fig. 9a), we can see that the left or right lung is divided into three parts near averagely. Therefore, we averagely divided the lung-tissue displacement into six parts (Fig. 9b) with a few tolerated biases. The average displacement of each part was qualified for prediction purposes.

The displacement of Surface data 2 was measured from the changed values of the markers instead of computing via the registration algorithm. Two surface displacements (end of inspiration and small inspiration phase; end of inspiration and end of expiration) were computed, as shown in Fig. 10.

2.6. Design and fine-tuning of neural network

The configuration of layers and nodes is likely to consider the *empirical findings* in literature or rely on *intuitions* from experience [21]. These are suitable starting points for finding better solutions to configure layers and nodes. However, they still require robust test harnesses and controlled experiments that include some basic strategies:

(1) *Random test*: tests random configurations for the number of layers and nodes in each layer.

(2) *Grid test*: a systematic search for the number of layers and nodes in each layer.

(3) *Heuristic test*: a structured search for the number of layers and the number of nodes in each layer according to specified search algorithms (e.g., genetic algorithm) [22].

(4) *Exhaustive test*: all combinations of the number of layers and the number of nodes in each layer are tested; this strategy is feasible for a simple network or dataset.

In this work, we selected the grid test to find the optimal number of *layers* and *nodes*. Moreover, we also considered other six hyper-parameters that affect the network convergence—*batch size, learning rate, training epochs, optimizer, loss function,* and *activation function*. Hence, we need to tune eight parameters and hyper-parameters for better convergence of neural networks.

In our neural network, the data x is propagated forward to obtain the predicted value

$$y_{pred} = f_w(x)$$

where $f_w(\bullet)$ denotes neural networks, such as multi-layer perceptron networks. Before prediction, the weight of a neural network should be updated until the convergence by minimizing the loss function, e.g., mean square error

$$Loss_{MSE} = \sum_i (p_i - q_i)^2$$

where p denotes probability distributions of output (observed real value). q represents probability distributions of expectation $(p, q \in (0, 1))$. i represents the index of training data in a batch. In training, the data should be divided into training data and validation data.

Dividing the data for training and validation purposes has a significant impact on the final network convergence. Currently, the k-fold cross-validation method better solves the problem caused by the data division. The loss value for training is shown below

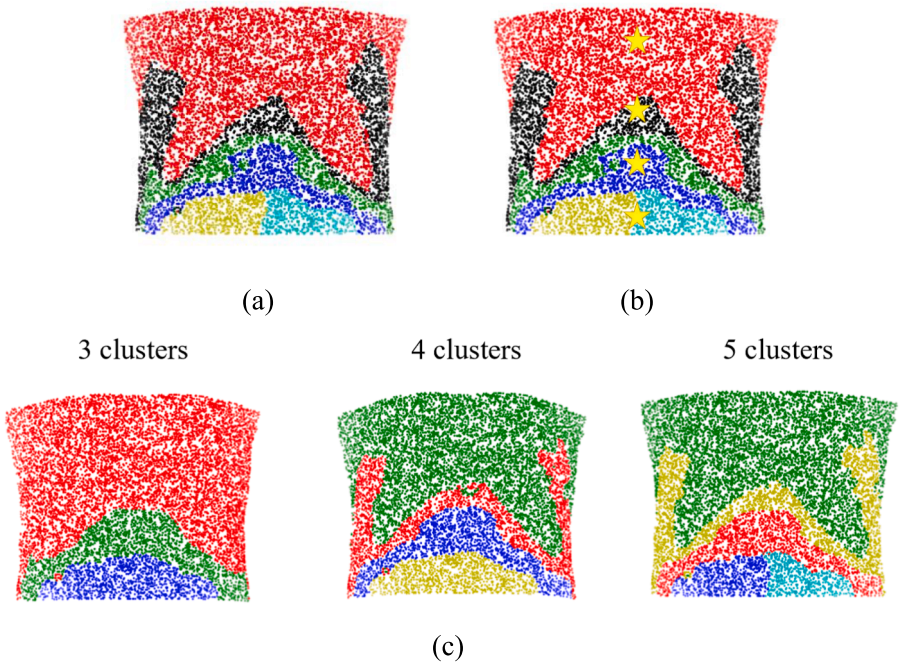


Fig. 7. (a) Clustering result based on Surface displacement 1, which divided the surface into six parts. (b) The final marker placement positions. (c) Different surface clustering results with 3, 4, and 5 clustering centers respectively. It is easy to see that there is little difference in the clustering results with 4 and 5 clustering centers. The only difference is at the bottom of clustering result near the belly where one class is divided into two classes, because the surface in these two classes slightly moves towards different directions. There is little difference in the movement magnitude between these two classes according to human experience.

Table 1
Comparisons of initial and optimized data collection plans.

Items of the data collection	Initial data collection plan	Optimized data collection plan
Number of images to collect	2 CT images	3 CT images
Points of respiration cycle where data is collected	End of inspiration Large inspiration phase	End of inspiration Small inspiration phase End of expiration
Data type	3D CT	3D CT
Were surface markers used?	No	Yes
Surface data	2 CT images	4 marker values

$$Loss_{k\_CV} = \frac{1}{k}\sum_{i=1}^{k} LOSS_{MSEi}$$

where i denotes the index of fold in the k-fold cross-validation. Hence, k weights are updated by

$$w_i^{new} = w_i - \eta \bullet \frac{\partial Loss_{k_{CV}}}{\partial w_i}, i \in 1, 2, ..k$$

where η denotes the learning rate and i denotes the index of fold in the k-fold cross-validation.

To compare with traditional machine learning approaches, we also tested the linear regression method and two SVR methods in our dataset. Displacement prediction based on the linear method (least-square regression) is achieved by fitting the parameters c_1 and c_1 where

$$y_{linear} = c_1 \bullet x + c_2$$

Displacement prediction based on SVR was achieved by minimizing

$$\min_{w,b} \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + C\sum_{i=1}^{N} \max(0, \left|y_i - \left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\phi}(x_i) + b\right)\right| - \varepsilon)$$

where $\boldsymbol{w}$ and b are the weight vector and bias of the SVR. N is the number of training data (training instance). i denotes the index of training instance. C denotes the penalty to a training instance when this training instance is misclassified or within the margin boundary. SVR adopts the epsilon-insensitive loss where errors of less than ε are ignored.

3. Results and discussion

3.1. Data collection and visualization

The validation experiment of the lung-tissue displacement prediction

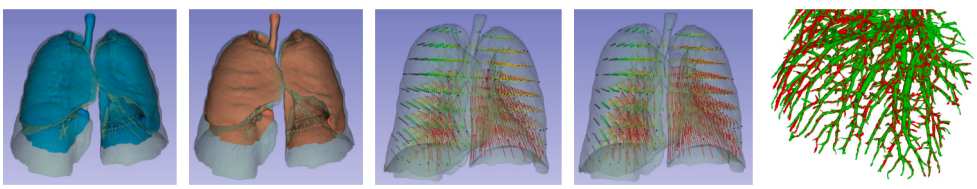


Fig. 8. Results after processing lung data 2 (3 phases). (a) Overlapping lung contours at the end of inspiration (green) and small inspiration phases (blue). (b) Overlapping lung contours at the end of inspiration (green) and end of expiration (red). (c) Lung-tissue displacements 2 at the end of the inspiration-small inspiration phase. (d) Lung-tissue displacements 2 at the end of inspiration-end of expiration. (e) Overlapping of vessels by the registration at the end of inspiration (green)-end of expiration (red).

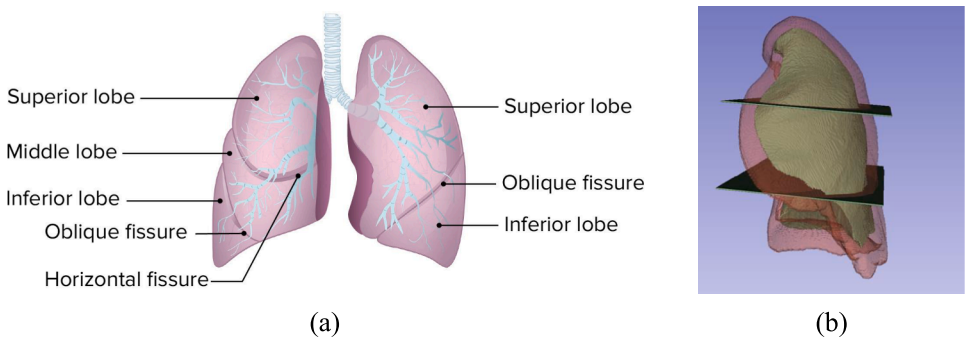


Fig. 9. (a) Anatomic representation of the lung and (b) fast method for lung division.

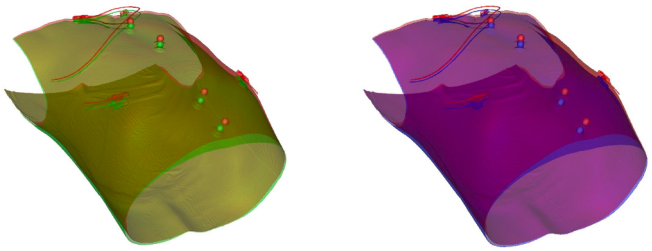


Fig. 10. Results after processing surface data 2 (three phases) with data obtained by value changes of surface markers. (a) Overlapping surface contours at the end of inspiration (red) and small inspiration phase (green). (b) Overlapping surface contours at the end of inspiration (red) and end of expiration (blue).

model was based on multi-phase pulmonary CT imaging data from 199 patient volunteers with 105 males and 94 females (Fig. 11a), across all ages (Fig. 11b).

After processing lung data 2, the displacements of six lung areas were obtained. The distributions of each area are presented in Fig. 11c. The lower lung area (left and right) near the inferior belly has the largest displacements, then the middle lung areas near the upper belly. The upper lung areas near the chest have the smallest displacements as the ribs constrain their movement. We also present the relationship between marker displacement and left lower lung area (Fig. 11d) to validate the efficacy of the following neural network models in non-linear lung-tissue displacement prediction.

3.2. The fine-tuning of neural networks

The design and fine-tuning of neural networks are based on eight parameters and hyper-parameters: number of layers, number of hidden units, batch size, learning rate, training epochs, optimizer, loss function, and activation function. The fine-tuning range of each parameter is shown in Table 2.

Detailed fine-tuning results are shown in Fig. 12, with the best-performing values of each parameter highlighted. The networks with fewer layers (2 and 3) outperform those with more layers (4 and 5) because a complex network with simple input may result in underfitting. Moreover, the results of 4 and 5 layers have higher deviations than that of 2 and 3 layers. The results of the 2 and 3 layers show little difference, and we selected three layers for training. A network with 16 hidden units performed best of all choices. Batch size has less impact on the training, and a batch of 50 slightly outperforms the rest of the options. Optimum learning rate is essential in network convergence, and learning rate 1e-4 performs best. A training epoch of 1,000 performs the best because of its smaller loss variance. The adam optimizer outperforms the others, but its advantage over SGD and RMSprop is relatively small. Loss functions MSE and MAE perform similarly in convergence, with the performance of MAE being slightly better. The activation function has a significant impact on network convergence. ReLU and LeakyReLU perform nearly the same as each other, but ReLU slightly outperforms LeakyReLU in convergence.

3.3. The lung-tissue displacement prediction and comparisons

We applied 5-fold cross-validation in the training process. After fine-tuning, the network with optimal parameters and hyper-parameters was obtained. The weight of the neural network was also saved for prediction (regression) purposes.

R-Squared, the coefficient of determination [23], is a statistical measure in a regression model that determines the proportion of variance in the dependent variable that the independent variable can explain. In other words, R-squared shows how well the data fit the regression model.

The formula for calculating R-squared is:

$$R^2 = \frac{SS_{regression}}{SS_{total}}$$

where $SS_{regression}$ is the sum of squares due to regression and SS_{total} is the total sum of squares.

We compared our proposed deep neural network with three traditional machine learning approaches, including the linear algorithm [11], SVR with linear core [16], and SVR with radial basis function (RBF) core [16]. As shown in Table 3, our proposed neural network outperforms these three machine learning approaches, where k denotes the kth cross-validation.

With respect to the time complexity, the time complexity of our neural network is computed by $\mathcal{O}(nt^*(ij+jk))$ where t is the number of epochs. n is the number of training instances or training samples. i, j and k are the numbers of hidden units in the first layer, second layer, and third layer respectively. Given the fine-tuning results in Table 2, the time complexity of our neural network is $\mathcal{O}(50^*n^*(16^*16 + 16^*16))$. The time complexity of SVR is $\mathcal{O}(nd)$ for linear kernel and $\mathcal{O}(n^2d)$ for RBF kernel where n is the number of training samples, and d is the dimension of training samples. The dimension of samples is 2 in this paper, therefore the time complexity of SVR is $\mathcal{O}(2n)$ for linear kernel or $\mathcal{O}(2n^2)$ for RBF kernel. The complexity of least-square regression is $\mathcal{O}(d^2(n+d))$ where n is the number of training samples, and d is the dimension of samples. Therefore, the time complexity of least-square regression is $\mathcal{O}(4(n+2))$.

In the experiment, the training of our neural network takes 2 min to converge in Intel core i7. The training of our neural network is fast. Hence, in the time complexity, the advantage of least-square regression and SVR with linear kernel over our neural network can be ignored.

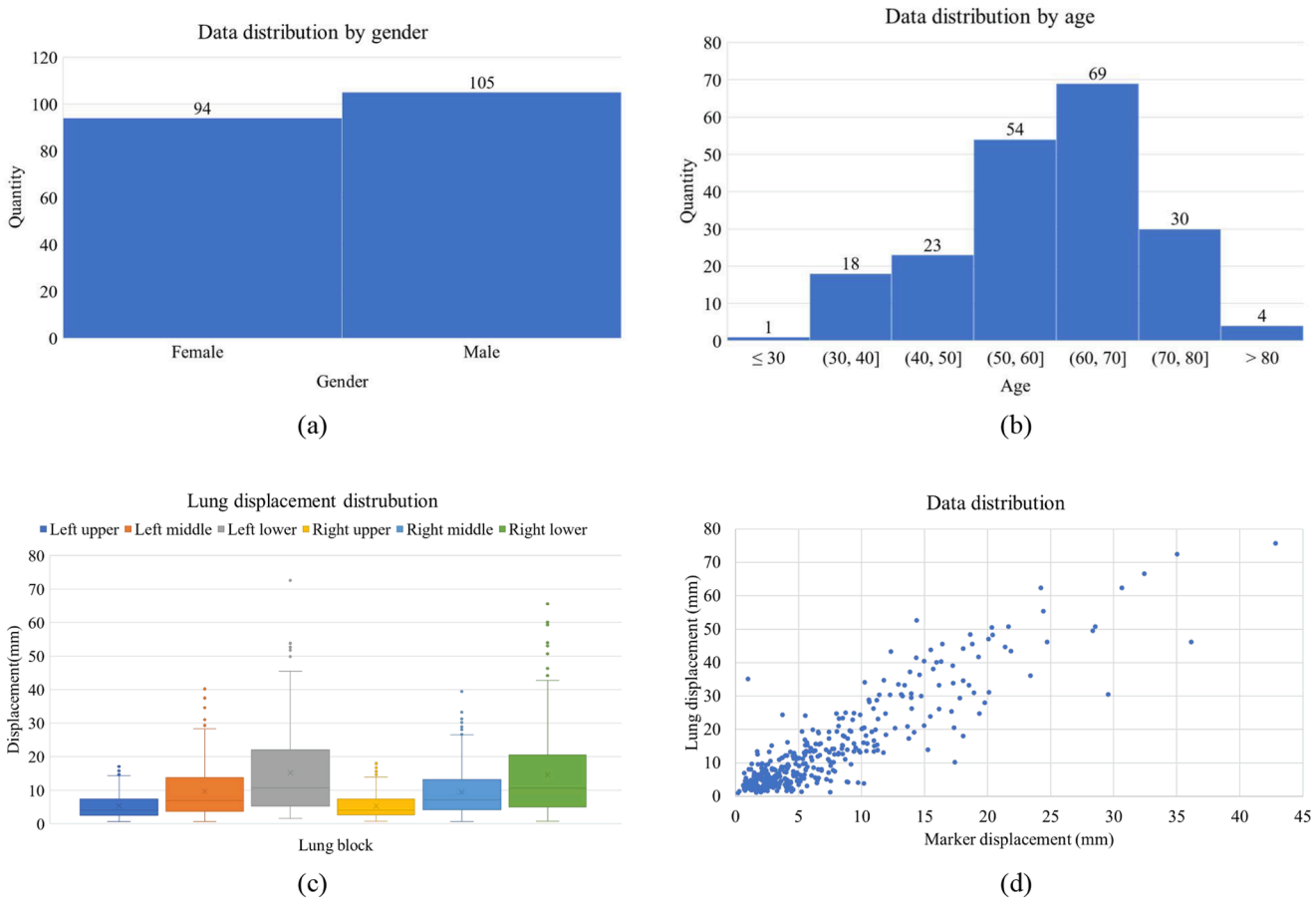


Fig. 11. Data distribution in our pulmonary CT imaging dataset. (a) The number of patients classified by gender. (b) The number of patients classified by age. (c) Displacement distribution of the six lung areas. (d) Relationship between lung-tissue displacement (the inferior lung block) and marker displacement.

Table 2
The fine-tuning parameters, hyper-parameters, and their optional values in the experiments, with the final selected value in Bold.

Parameters and hyper-parameters	Considered values
Number of layers	2, **3**, 4, 5
Number of hidden units	8, **16**, 32, 64
Batch size	25, **50**, 100, 200
Learning rate	1e-2, 1e-3, **1e-4**, 1e-5
Training epoch	100, 250, 500, **1000**
Optimizer	SGD, RMSprop, adagrad, **adam**
Loss function	hinge, MSE, **MAE**, MAPE
Activation function	sigmoid, tanh, **ReLU**, leakyReLU

4. Discussion and conclusion

The weakness of this paper is that collected data currently exhibits more linearity which brings bias to the experiments, given the result of data visualization in Fig. 11d. Moreover, it is easy to see that more collected data is distributed at the bottom-left, and less data is distributed at the upper-right. This means less CT images are collected at the end of expiration. This results in the marginal advantage of neural network over classical methods in the prediction. With more data collected, we believe that data will exhibit more nonlinearity because human respiration curve is nonlinear. The neural network is good at handling the prediction in nonlinear cases. The prediction accuracy of the neural network still has the space to improve. Other ways to improve the prediction accuracy may consider the time-sequential property of respiration process. Time-related features should be added to the features of training data (the displacements of lung tissue and body surface) to indicate which respiration phase the training data is collected, therefore improving the feature quality. Sometimes, the same surface displacement does not correspond to the same lung-tissue displacement if the displacements are computed using the CT images collected from different respiration phases. This results from the nonlinearity of the respiration. Future works will focus on collecting more data with higher displacements to further improve the prediction accuracy of neural networks. More prediction algorithms like the long-short term memory (LSTM) [24] and Bayesian estimation [25] may be examined if the time-sequential property of the respiration is considered.

Percutaneous puncture is essential for the interventional diagnosis and treatment of lung tumors. The success of this operation is dependent on accurate needle placement. However, respiratory movements may reduce its accuracy and efficiency. To improve the success of lung percutaneous puncture, it is essential to investigate how lung tissues move during respiration.

This paper systematically developed a deep learning model for lung-tissue displacement prediction using multi-phase pulmonary CT Images. First, lung and surface CT data from two phases of respiration are processed to investigate lung and body-surface movements. The data in this stage provides a new, optimized data collection plan. Second, new lung and surface CT data are collected from three phases of respiration (the end of expiration, the small inspiration phase, and the end of inspiration) and then processed. The data in this stage provides a training dataset for a deep neural network model to predict lung-tissue displacement. Third, several network training experiments are performed to determine the optimal parameters and hyper-parameters. The weights of the network with optimal parameters and hyper-parameters are saved for prediction purposes. The validation experiment (the

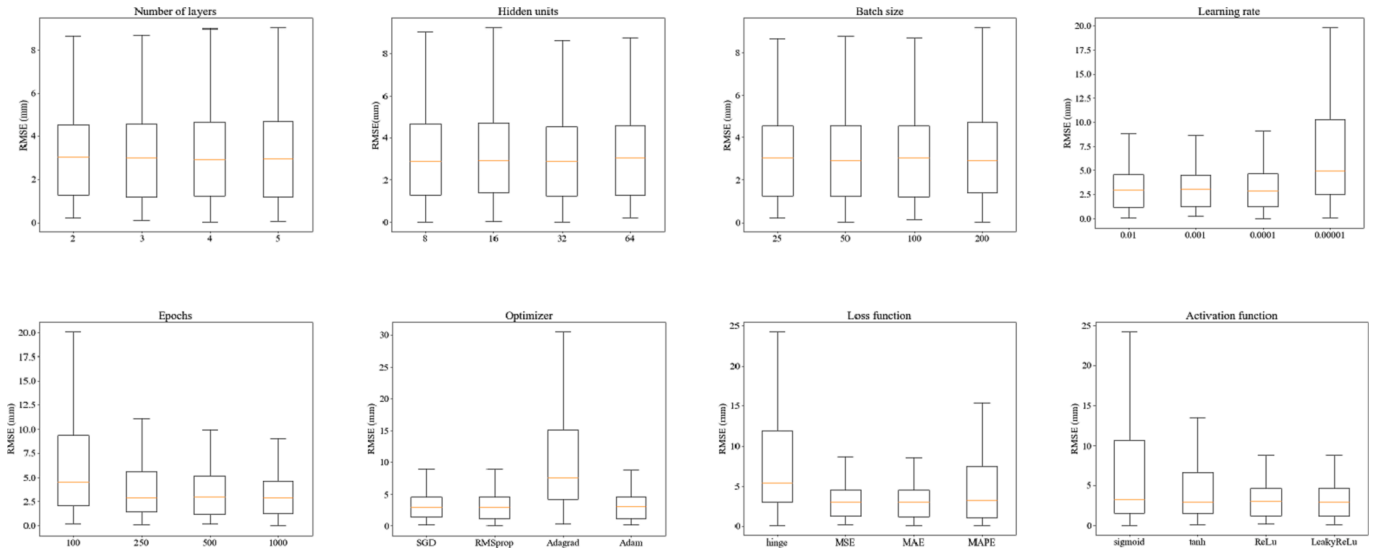


Fig. 12. Fine-tuning results of eight parameters and hyper-parameters in our proposed deep neural network.

Table 3
The R^2 value of different approaches. In SVR, the penalty $C = 1.0$. Stopping tolerance is set to 0.001, and epsilon $\varepsilon = 0.1$.

Approach	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	Average
Linear	0.67	0.80	0.80	0.85	0.84	0.792
SVR (linear)	0.66	0.81	0.80	0.85	0.84	0.792
SVR (RBF)	0.58	0.71	0.63	0.69	0.74	0.672
Our Neural network	**0.68**	**0.81**	**0.80**	**0.85**	**0.85**	**0.798**

displacement prediction of the left lower lung) demonstrates that the accuracy of our designed network outperforms that of three traditional machine learning approaches including the linear model, linear support vector regression model, and support vector regression model with radial basis function. In conclusion, the clustering results make a set of our designed methods reliable. Our neural network method can predict lung-tissue displacement based on body-surface displacement during respiration, potentially improving the success of lung puncture procedures.

Funding source

This study was supported by the Project of the Educational Commission of Guangdong Province of China (No. 2022ZDJS113) and the School-Enterprise Cooperation Fund provided by Wuerzburg Dynamics Inc. to the Weiding Joint Laboratory of Medical Artificial Intelligence, Shenzhen Technology University.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

References

[1] R.L. Siegel, et al., Cancer statistics, 2022. CA: A Cancer J. Clin., 2022. **72**(1).
[2] J. Song, et al., Electromagnetic navigation bronchoscopy-guided dye marking for localization of pulmonary nodules, Ann. Thorac. Surg. (2021).
[3] G.B. Murrmann, F.H.M. Van Vollenhoven, L. Moodley, Approach to a solid solitary pulmonary nodule in two different settings—"Common is common, rare is rare", J. Thorac. Dis. 6 (3) (2014) 237–248.
[4] D. Sortini, et al., Thoracoscopic localization techniques for patients with solitary pulmonary nodule and history of malignancy, Ann. Thorac. Surg. 79 (1) (2005) 258–262.
[5] B. Chen, et al., Design and validation of a novel MR-compatible sensor for respiratory motion modeling and correction, IEEE Trans. Biomed. Eng. (2016) 123–133.
[6] F. Büther, et al., Clinical evaluation of a data-driven respiratory gating algorithm for whole-body PET with continuous bed motion, J. Nucl. Med. 61 (10) (2020) 1520–1527.
[7] M. Zaitsev, J. Maclaren, M. Herbst, Motion artifacts in MRI: A complex problem with many partial solutions, J. Magn. Reson. Imaging 42 (4) (2015).
[8] J.R. Mcclelland, et al., Respiratory motion models: A review, Med. Image Anal. 17 (1) (2013) 19–42.
[9] L. Zheng, et al., A Novel Respiratory Follow-Up Robotic System for Thoracic-Abdominal Puncture, IEEE Trans. Ind. Electron. 68 (3) (2021) 2368–2378.
[10] Y. Ma, et al., Target localization during respiration motion based on LSTM: A pilot study on robotic puncture system, Int. J. Med. Rob. Comput. Assisted Surg. 11 (2021).
[11] F. Ernst, et al., Evaluating and comparing algorithms for respiratory motion prediction, Phys. Med. Biol. 58 (11) (2013) 3911–3929.
[12] F. Ernst, et al., Correlation between external and internal respiratory motion: a validation study, Int. J. Comput. Assist. Radiol. Surg. 7 (3) (2012) 483–492.
[13] J.R. McClelland, et al., A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy, Med. Phys. 33 (9) (2006) 3348–3358.
[14] U. Wijenayake, S.Y. Park, Real-Time External Respiratory Motion Measuring Technique Using an RGB-D Camera and Principal Component Analysis, Sensors 17 (8) (2017).
[15] A.E. Torshabi, et al., Targeting Accuracy in Real-time Tumor Tracking via External Surrogates: A Comparative Study, Technol. Cancer Res. Treat. 9 (6) (2010) 551–561.
[16] L. Lei, et al., Automatic registration and precise tumour localization method for robot-assisted puncture procedure under inconsistent breath-holding conditions, The International Journal of Medical Robotics and Computer Assisted Surgery 17 (6) (2021) e2319.
[17] M. Isaksson, J. Jalden, M.J. Murphy, On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications, Med. Phys. 32 (12) (2005) 3801–3809.
[18] N. Gerber, et al., High-accuracy patient-to-image registration for the facilitation of image-guided robotic microsurgery on the head, IEEE Trans. Biomed. Eng. 60 (4) (2013) 960–968.
[19] Q. Lin, et al., Real-time automatic registration in optical surgical navigation, INFRARED PHYSICS AND TECHNOLOGY (2016).
[20] H. Dang, et al., Robust methods for automatic image-to-world registration in cone-beam CT interventional guidance, Med. Phys. 39 (10) (2012) 6484–6498.
[21] Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning*. MIT Press.
[22] Stathakis and D.,, How many hidden layers and nodes? Int. J. Remote Sens. 30 (8) (2009) 2133–2147.
[23] Di Bucchianico, A., *Coefficient of Determination (R2)*, in *Encyclopedia of Statistics in Quality and Reliability*. 2007.
[24] Y. Yong, S. Xiaosheng, H. Changhua, Z. Jianxun, A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (7) (2019) 1235–1270.
[25] C. Zhou, C. Wang, H. Hassan, H. Shah, B. Huang, P. Fränti, Bayesian inference for data-efficient, explainable, and safe robotic motion planning, A Review. arXiv: 2307.08024 (2023) 1–33.