Received: 12 March 2020 | Revised: 10 November 2020 | Accepted: 10 November 2020

DOI: 10.1002/rcs.2247

ORIGINAL ARTICLE

The International Journal of Medical Robotics and Computer Assisted Surgery

WILEY

Target localization during respiration motion based on LSTM: A pilot study on robotic puncture system

Yuxiang Ma[1,2] | Zhikai Yang[1,2] | Wei Wu[1,2] | Hongzhi Xie[3] | Lixu Gu[1,2]

[1]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

[2]Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China

[3]Department of Cardiology, Peking Union Medical College Hospital, Beijing, China

Correspondence

Lixu Gu, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China.
Email: gulixu@sjtu.edu.cn

Funding information

Beijing Municipal Natural Science Foundation, Grant/Award Number: L192006; the National Key research and development program, Grant/Award Number: 2016YFC0106200; the 863 national research fund, Grant/Award Number: 2015AA043203; the funding from IMR of SJTU

Abstract

Background: In the needle biopsy, the respiratory motion causes the displacement of thoracic–abdominal soft tissues, which brings great difficulty to accurate localization. Based on internal target motion and external marker motion, the existing methods need to establish a correlation model or a prediction model to compensate the respiratory movement, which can hardly achieve required accuracy in clinic use due to the complexity of the internal tissue motion.

Methods: In order to improve the tracking accuracy and reduce the number of models, we propose a framework for target localization based on long short-term memory (LSTM) method. Combined with the correlation model and the prediction model by using LSTM, we adopted the principal component of time-series features of external surrogate signals to predict the trajectory of the internal tumour target. Additionally, based on the electromagnetic tracking system and Universal Robots 3 robotic arm, we applied the proposed approach to a prototype of robotic puncture system for real-time tumour tracking.

Results: To verify the proposed method, experiments on both public datasets and customized motion phantom for respiratory simulation were performed. In the public dataset study, an average mean absolute error, and an average root-mean-square error of predictive results of 0.44 and 0.58 mm were achieved, respectively. In the motion phantom study, an average root mean square of puncturing error resulted in 0.65 mm.

Conclusion: The experimental results demonstrate the proposed method improves the accuracy of target localization during respiratory movement and appeals the potentials applying to clinical application.

KEYWORDS

respiratory motion prediction, robotic puncture system, tumour tracking, long short-term memory

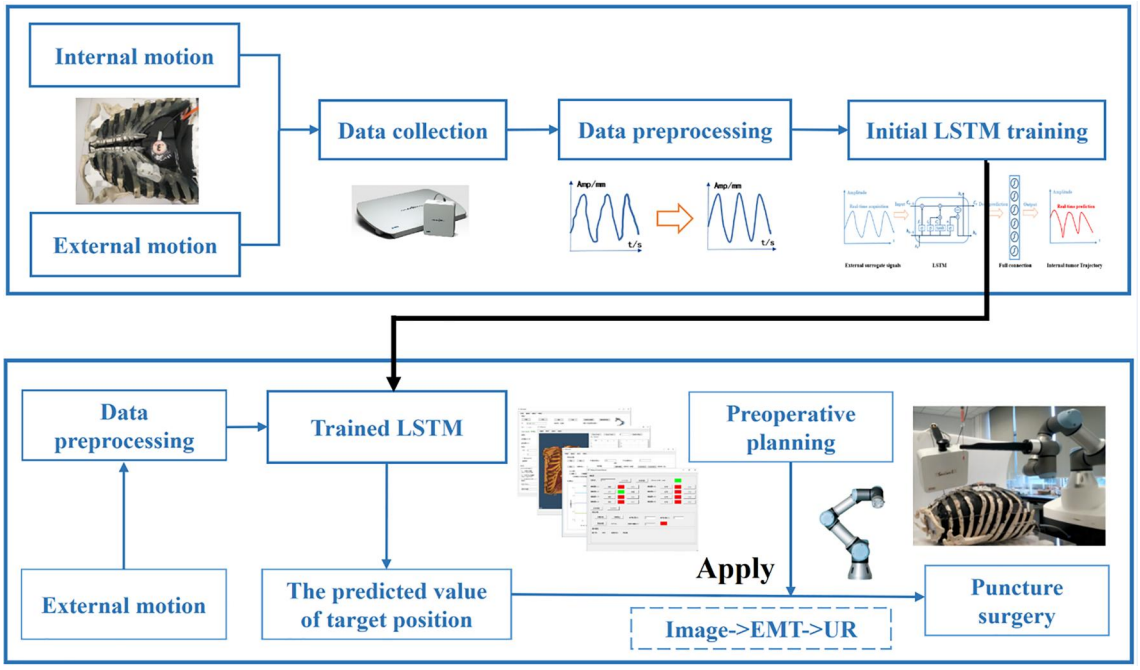

FIGURE 1 The proposed framework for target localization based on LSTM in robotic puncture system. EMT, electromagnetic tracking system; LSTM, long short-term memory

1 | INTRODUCTION

Recently, lung cancer has become the leading cancer killer worldwide, accounting for the largest number of deaths.[1] Due to a characteristically long latency period, it is difficult to be detected at an early stage. The common diagnosis method is CT-guided percutaneous biopsy. However, influenced by respiratory motion, patients' organs and tumours have complex three-dimensional (3D) motions, and the maximum range of motion can be up to 4.0 cm.[2] In order to accurately locate the internal lesion, surgeons require taking intraoperative CT verification multiple times. However, the increased radiation dose exerts a detrimental effect on patient's health. Thus, efforts addressed on respiratory compensation and respiratory tracking to improve the accuracy of target localization. Lee et al. developed a respiratory motion reduction device,[3] but it is difficult to achieve precise duplicate positioning in the course of surgery. Based on 4D CT, it has led to major advances in the respiratory movement modelling.[4] Unfortunately, the radiation dose is extremely high. With the development of synchronous respiratory tracking technology in the radiotherapy, the Cyberknife,[5,6] which is a direct soft tissue tracking method, improves the accuracy of tumour localization. It establishes a correlation motion model between internal target recorded by a bi-planar x-ray system and external optical markers, such as a linear or polynomial model. In order to further improve tracking accuracy, quite a few efforts have been addressed.

In the studies of correlation model, Floris et al. placed gold fiducials into the swine's livers and multiple LEDs on the swine's chest. They established a model based on ε-SVR (support vector regression) to investigate the correlation of the external signal to the motion of the liver.[7] After that, Floris kept on using the optical markers and the ultrasound station to acquire the external data and internal measured data of the human body, respectively. He correlated the external data to the principal component of the internal data, and resulted a root mean square (RMS) of 1.29 mm.[8] Matteo et al. established an artificial neural network model for each of three markers on patients' skin to estimate the tumour motion.[9] The model was checked and updated periodically when the prediction error increased. Dou et al. took the sensor noise and the model error into consideration, and proposed a correlation model based on unscented transformation (UT) to reduce the uncertainty in the modelling process.[10]

In the studies of prediction model, Hong et al. applied the Kalman filter method to make an optimal estimation of the input data based on the state transition equations.[11] In order to attain regularity of breathing signals, Sun et al. proposed a neural network using adaptive augmentation and multilayer perceptions.[12] Wang et al. applied a seven-layer bidirectional long short-term memory (bi-LSTM) to predict the respiration motion for external surrogate signals.[13]

Based on internal target motion and external marker motion, the existing methods need to establish a correlation model or a prediction model to compensate the respiratory movement, which can hardly achieve required accuracy in clinic use due to the complexity of the internal tissue motion. A prediction model predicts the next position for external data. And a correlation model correlates the external data to the internal data at the same time point. For real-time free-breath target localization, the correlation model plays a more significant role than a prediction model.[12] Considering the characteristics of time series of internal/external motion, we propose a framework that combines the above two models by using LSTM. Our method adopted historical external surrogate signals to predict the trajectory of the internal tumour target at the next moment directly to improve the tracking accuracy. Then, based on the electromagnetic tracking system

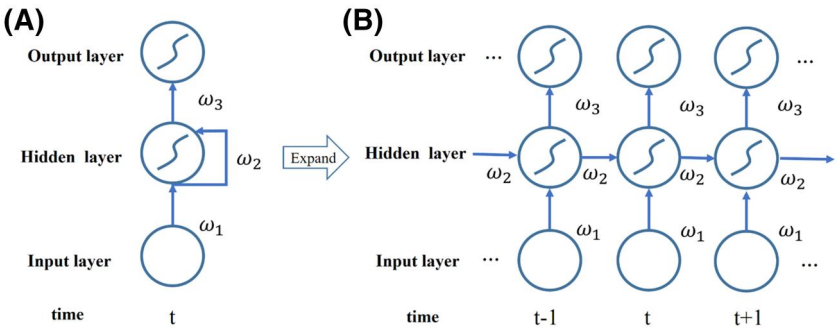

FIGURE 2 The structure diagram of recurrent neural network. The ω_1 is input-to-hidden layer weight matrix. The ω_2 is the weight matrix between hidden layers. And the ω_3 is hidden-to-output layer weight matrix

(EMT) and Universal Robots 3 and (UR3) robot arm, we applied the proposed approach to a prototype of the robotic puncture system, which functions of image reconstruction, puncture path planning, robot controlling, multi-coordinate registration and real-time free-breath tumour tracking. The puncture robot can automatically adjust the needle insertion pose and subsequently target the tumor's movement in an improved accuracy during the free regular respiratory guided by our proposed method as shown in Figure 1.

2 | MATERIALS AND METHODS

2.1 | The prediction method based on LSTM

Since respiration is a periodic motion, external surrogate signals and the internal tumour target, respectively, constitute time-series data. The recurrent neural networks (RNNs) have been proven to be a powerful tool in solving the problem of time series.[14] Unlike traditional algorithms in machine learning, it does not require manually extracting features from the dataset, but adopts the back propagation mechanism to adjust internal parameters by itself. Therefore, RNN has been widely applied in many fields, such as glucose prediction[15] and vehicle trajectory prediction.[16]

RNN adopts a recursive method by using the network's own recurrent structure to record historical data. The hidden layers of the RNN are interconnected, and the weights of each node at different times are shared. Therefore, the output of the current network node depends on the current input and the output of the hidden layer at the previous moment. Figure 2a shows the internal structure of RNN, and Figure 2b shows the structure of RNN unfolded on the timeline. With the continuous development of RNN, other forms such as BiRNN[17] and LSTM[18,19] are gradually derived.

Bidirectional recurrent neural network (Bi-RNN) is an enhanced version of RNN. Since in many practical tasks, such as semantic analysis, the understanding of the current word does not depend on only the word that appeared before, but also closely related to the word that follows it. Bi-RNN appears to ensure that the context information of the current network nodes can be taken into account when processing data. As shown in Figure 3, Bi-RNN has the superposition hidden layer of the forward RNN layer and the backward RNN layer based on the original RNN. The common characteristics of historical information and future information are applied together at the output layer.

However, with the increase of the distance between the relevant information and the current prediction node, the traditional RNN will cause the gradient disappearance or the gradient explosion in the training process due to the increase in the number of layers in the network. As shown in Figure 2, the chain structure of the RNN has a repetitive and monotonous neural network module, such as a simple tangent layer. Therefore, LSTM adopts the unit to replace the hidden nodes in the traditional RNN, so as to analyse the data features with context dependence and solve the problem of updating failure after gradient transmission, as shown in Figure 4.

For effective regulation, an LSTM unit includes a cell and three gates (an input gate, an output gate and a forget gate). When the data is transferred to an LSTM unit, it is processed according to the following algorithm.

2.2 | The algorithm of an LSTM unit

As shown in Figure 4, x_t、 C_t and h_t, respectively, represent the vector of input, cell state and output gate's activation at time t.

2.2.1 | Step 1

The forget gate determines the retention of the vector at the last moment. The vector f_t of the forget gate can be expressed as

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)\sigma \in (0, 1) \quad (1)$$

where $[h_{t-1},\ x_t]$ is a vector from the concatenation of h_{t-1} and x_t, W_f and b_f are the corresponding weight matrices and bias vector parameters, and σ is the sigmoid function.

2.2.2 | Step 2

The input gate is used to decide which information is going to be updated, which is affected by the current input vector and the previous output vector. Assume that W_i and b_i are the corresponding weight matrices and bias vector parameters, the vector i_t of the input gate can be expressed as

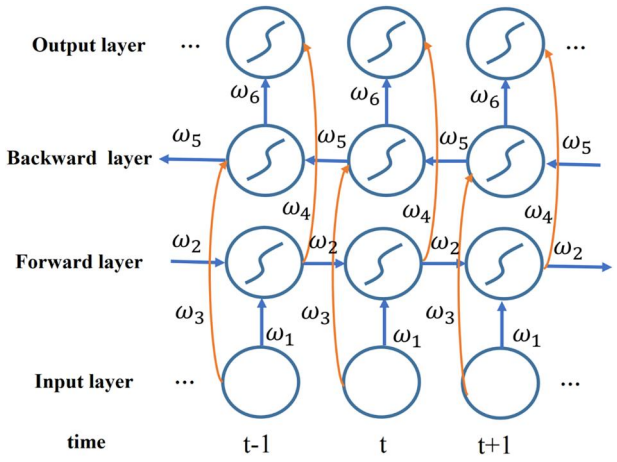

FIGURE 3 The structure diagram of bidirectional recurrent neural network. The ω_1, ω_3, ω_4 and ω_6 represent the weight matrix of input-to-forward layer, input-to-backward layer, forward-to-output layer and backward-to-output layer, respectively. And the ω_2 and ω_5 represent the weight matrix of forward layers and backward layers, respectively

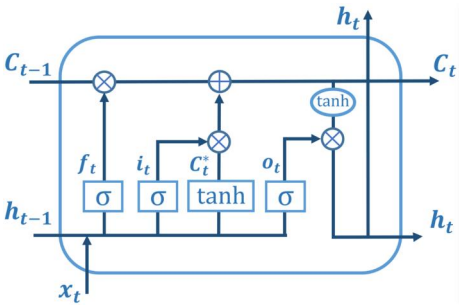

FIGURE 4 The architecture of a long short-term memory unit. The f, i and o are, respectively, the forget gate, input gate and output gate. And x, C*, C and h, respectively, represent the vector of input, candidate cell state, cell state and output gate's activation

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)\sigma \in (0, 1) \tag{2}$$

2.2.3 | Step 3

The candidate cell state C_t^* is calculated by tanh for h_{t-1} and x_t. W_C and b_C are the corresponding weight matrices and bias vector parameters:

$$C_t^* = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

2.2.4 | Step 4

The current cell state is determined by the old state C_{t-1} multiplied by f_t, and C_t^* multiplied by i_t:

$$C_t = f_t \times C_{t-1} + i_t \times C_t^* \tag{4}$$

where $\times$ is the pointwise multiplication.

2.2.5 | Step 5

The vector o_t of the output gate is determined by h_{t-1} and x_t, and W_o b_o are the corresponding weight matrices and bias vector parameters:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)\sigma \in (0, 1) \tag{5}$$

2.2.6 | Step 6

According to formulae (4) and (5), we can obtain the output vector of an LSTM unit by the output gate and the cell state:

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

Because RNN has good characteristics of time-series prediction, we propose a framework based on LSTM for respiratory motion tracking. We adopt the memory capability of LSTM for data analysis and processing to extract the complex mapping features of time series data.

Assume that external surrogate signals and the corresponding trajectory of the internal tumour target is $\boldsymbol{X}_{i:j} = [X_i, X_{i+1}, \ldots, X_{j-1}, X_j]^T$, $\boldsymbol{Y}_{i:j} = [Y_i, Y_{i+1}, \ldots, Y_{j-1}, Y_j]^T$, respectively, from time i to j, where $X_i = [X_i^{(0)}, X_i^{(1)}, X_i^{(2)}]^T$, $Y_i = [Y_i^{(0)}, Y_i^{(1)}, Y_i^{(2)}]^T$. In order to further explore the correlation relationship between them, we use principal component analysis[20] to reduce the dimensions of $\boldsymbol{X}_{i:j}$ and $\boldsymbol{Y}_{i:j}$, respectively, to 1D principal component $\boldsymbol{X}'_{i:j} = [X_i', X_{i+1}', \ldots, X_j']^T$ and $\boldsymbol{Y}'_{i:j} = [Y_i', Y_{i+1}', \ldots, Y_j']^T$. Assume that the size of a sliding window for the time-series prediction is k ($k < j - i + 1$), the prediction algorithm based on LSTM is shown in Figure 5. The model structure is mainly composed of LSTM network layer and fully connected layer. Taking example for predicting Y'_{i+k} from $X'_{i:i+k-1}$, the input of the algorithm is $\boldsymbol{X}^*_{i:j-1} = [X'_{i:i+k-1}, X'_{i+1:i+k}, \ldots, X'_{j-k:j-1}]$, and the expected output is $\boldsymbol{Y}^*_{i+k:j}$ $[Y'_{i+k}, Y'_{i+k+1}, \ldots, Y_j']^T$. After continuous iterative training, the prediction algorithm has the characteristics of accurately predicting for trajectory of the internal tumour target gradually.

2.3 | Robotic puncture system

2.3.1 | EMT and UR3

EMT is a spatial measurement system, which is used to track surgical tools with the sensor coils intraoperatively. We adopted a planar field generator, NDI aurora system[21] which is from Northern Digital Inc., Canada.

UR3 is a well-known 6-DoF (degree-of-freedom) robotic manipulator manufactured by Universal Robots company in Denmark. It has a series of advantages, such as easy programming, fast set-up, flexible deployment, collaborative and safe.[22] Besides, the accuracy of its end effector is ±0.1 mm, which can meet the

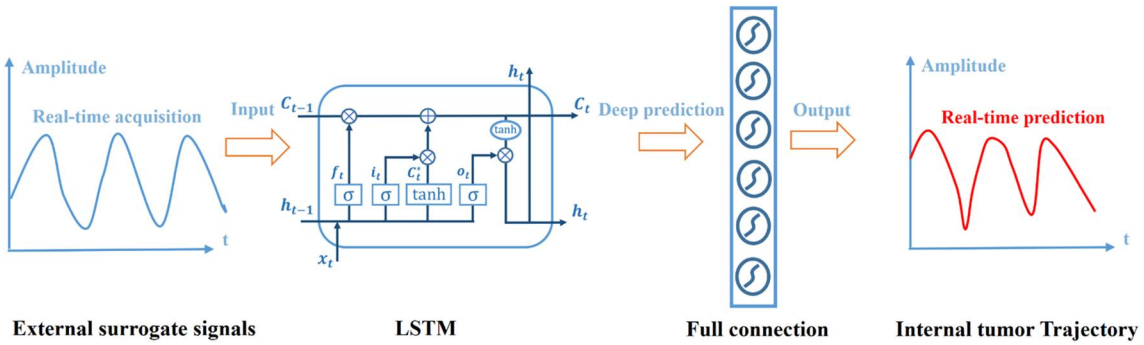

FIGURE 5 The prediction algorithm based on long short-term memory

requirements of surgery. Using the script provided by its official manual, we can easily control UR3. Combined with the puncture surgery scene, we processed on the basis of the UR3. We adopted polysulphide (thermoplastic resin) to make a lightweight, equal-torque and partially hollow support rod, which was fixed to the flange of the end effector in Figure 6. It can provide a fixed place for the puncture needle and is beneficial to prevent metal parts inside the UR3 from interfering with the magnetic field area. A 5-DoF sensor was embedded in the tip of the needle and calibrated to the tool centre point in UR3 for tracking by EMT.

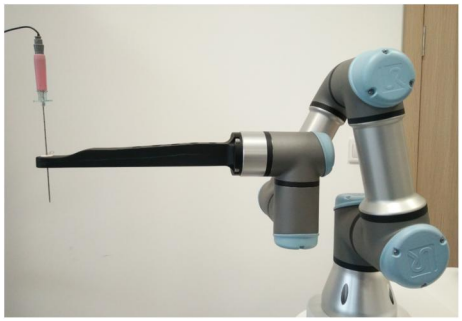

FIGURE 6 The extended end-effector of Universal Robots 3

2.3.2 | Multi-coordinate registration

In order to determine the pose of the robot for puncturing, we register image reconstruction coordinate, electromagnetic coordinate and UR coordinate using point-based registration.[23]

First, the image reconstruction coordinate and the electromagnetic coordinate are registered to get the spatial transformation matrix $T_{Mag\leftarrow CT}$. These landmarks were fixed on where respiratory movement is not significant to avoid the effect of respiration motion, such as the sides of the ribs. Similarly, the matrix $T_{UR\leftarrow Mag}$ is determined in electromagnetic coordinate and UR coordinate system. Because the phantom is placed in the best measuring volume of magnetic field, to avoid the influence of location and spatial extend, registration points around the model were collected by UR arm with a sensor attached in advance. Thus, the transform relationship from image coordinate to UR coordinate can be expressed as follows:

$$P_{UR} = T_{UR\leftarrow Mag} T_{Mag\leftarrow CT} P_{CT} \quad (7)$$

3 | EXPERIMENTAL RESULTS

To verify the performance of the proposed method, two kinds of experiments were performed. First, the proposed method was applied on public datasets and compared with existing methods. Then, with the developed robotic puncture system, simulation experiments were carried out on a customized respiratory motion simulation phantom. The software system is built in a visual studio 2015 with open-source visualization toolkits and libraries: vtk7.1 (vtk.org), Qt5.10 (www.qt.io) and keras2.2 (keras.io). And the computing devices were a workstation with Intel® Xeon(R)CPU E5-2620 v3, TITAN X Pascal GPU and a laptop with Intel® Core(TM)i5-9300H CPU, GTX 1650.

Mean absolute error (MAE) and root-mean-square error (RMSE) were introduced for evaluating the prediction method. RMS and TRE (target registration error) were introduced for the error measurement of coordinate registration.[24] RMS was introduced for assessing the accuracy of the target puncturing.

3.1 | The verification of the LSTM method

To evaluate the performance of the proposed LSTM prediction method, a public dataset[25,26] from the Institute of Robotics and Cognitive Systems, Lubeck University, Germany, which includes seven male patients aged between 23 and 30, was adopted. The dataset contains bimodal respiratory motion traces, one of which is an optical marker placed on the patient's chest, and the other is a vessel bifurcation recorded by 4D ultrasound with 17.5–21.3 Hz sampling. The recording length of each subject ranges from 6433 to 8169, with a total of 50 344. We removed outliers, intercepted the stable part and smoothed the data by Savitzky–Golay algorithm.[27] In order to eliminate the difference between external surrogate signals and the trajectory of the internal tumour target in the bimodal properties, we normalized the data with z-score to ensure that the

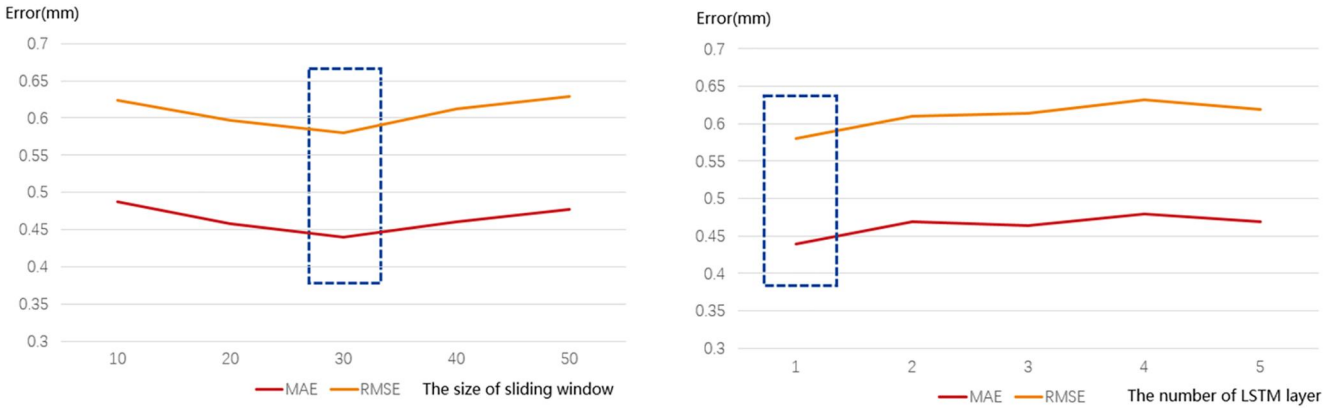

FIGURE 7 (A) The prediction performance under different sliding window sizes. (B) The prediction performance under different long short-term memory layers

TABLE 1 The prediction average error of the target for the test set

MAE (mm)	RMSE (mm)
0.44	0.58

Abbreviations: MAE, mean absolute error; RMSE, root-mean-square error.

gradient of LSTM moved towards the minimum value and improved the convergence speed of model training. Then, we got the length of the time series for each subject ranging from 3548 to 7862, with a total of 40 200. Due to the specificity of each sample, we adopted 60% of each patient dataset (87 respiratory cycles) as the training set, 20% (29 respiratory cycles) as the validation set, and 20% (29 respiratory cycles) as the test set. The basic super parameters of the network model were determined by the whole group training data as follows: the optimizer, the learning rate, the batch size and the epoch were Adam, 0.001, 512 and 100, respectively. Then, the prediction performance under different sliding window sizes and LSTM layers was analysed.

First, we set the size of different sliding window of LSTM, which was represented by a line chart in Figure 7a. The abscissa represented the size of sliding window from 10 to 50 in 10-step increments each time, and the ordinate represented the prediction accuracy of the target in the test set. With the increase of the size of sliding window, the MAE and RMSE showed a trend of first decreasing and then increasing. Therefore, the best prediction accuracy can be obtained by selecting a size of sliding window of 30 for historical data in LSTM.

Second, we set the number of different layers of LSTM, which was represented by a line chart in Figure 7b. The abscissa represented the number of LSTM layer from one to five and the ordinate represented the prediction accuracy of the target in the test set. With the increase of the number of layers, the structure of LSTM became more and more complex. However, the MAE and RMSE did not decrease accordingly, but had an increasing trend. Therefore, the best prediction accuracy can be obtained by selecting a one-layer LSTM.

As shown in Table 1, the average MAE achieved 0.44 mm when the average RMSE was 0.58 mm. As an example, we randomly selected the predicted and actual values of more than 900 consecutive pairs of the internal tumour target for a single subject, as shown in Figure 8a. The abscissa represented the number of data points, and the ordinate represented the amplitude of each point. The red sequence represented the true values, and the blue sequence represented the predicted values. It reveals that the target motion tracking method based on LSTM has better prediction accuracy.

We also compared with other RNN methods, such as BiLSTM,[28] GRU[29] and BiGRU.[30] On the premise that the best parameters obtained from the above discussion (one layer, the size of sliding window is 30) were selected, we performed on the same test set, and the results were shown in Figure 8b. The ordinate represented the prediction accuracy of the target in the test set. It reveals the LSTM is superior to other RNN methods in the effect of prediction accuracy, which achieves the best accuracy of prediction.

In order to further verify the accuracy of the method proposed in this paper, we compared with ε-SVR[8] proposed by Floris who provided the public datasets. Because of the evaluation of the predicted value of the principal component of the target, the RMS adopted by Floris in the experiment was the same as the RMSE adopted here. As shown in Table 2, it reveals that our method is significantly improved in accuracy and robustness.

3.2 | The verification of target puncturing

In order to further simulate the thoracic–abdominal movement of human respiration, we built a respiratory movement simulation phantom. It consisted of an air compressor, a solenoid valve, a throttle valve, a multi-channel USB relay and an improved human skeleton model (including two reservoir bags), as shown in Figure 9. The air compressor was responsible for supplying air to the respiratory movement phantom. The solenoid valve was mainly used to control the three states of inhalation, deflation and breath holding. The throttle valve was used to regulate the air flow speed of

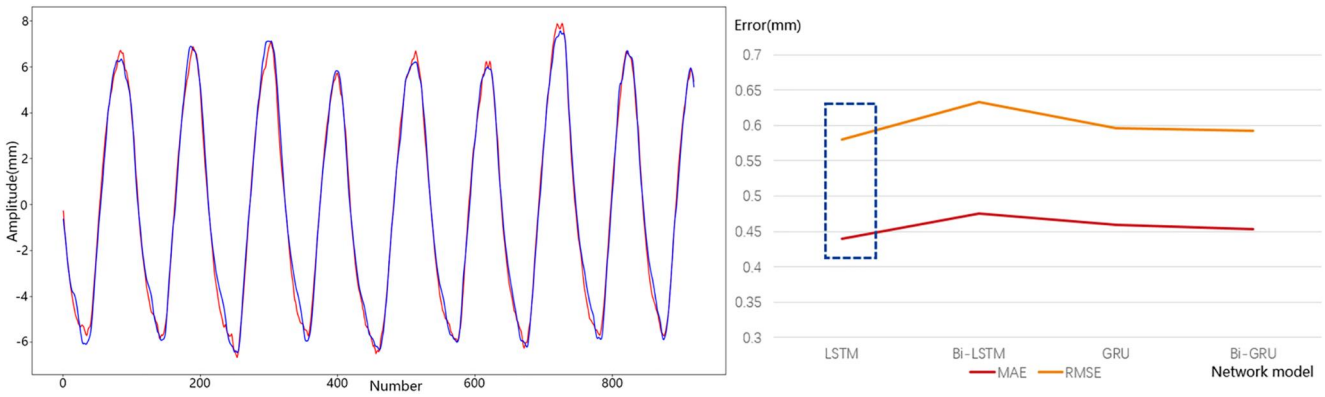

FIGURE 8 (A) Exemplifying predicted values and actual values of the internal tumour target (B) compared with other network methods

TABLE 2 Comparison of an existing method

Methods	MAE (mm)	RMSE (mm)
Floris ε-SVR	—	1.29
Our method	0.44	0.58

Abbreviations: MAE, mean absolute error; RMSE, root-mean-square error.

inhalation and exhalation. The multi-channel USB relay was used to simulate the normal respiratory movement of human body. Additionally, NDI authorized sensor markers were applied in the experiments. Four fiducial landmarks (M1–M4) were applied on the surface of the respiration motion model. An abdominal marker (M5) was used to record external surrogate signals. And another marker (M6) which is regarded as the target was fixed inside the phantom (the upper side of the reservoir bag). We adopted the position of M6 and that of the sensor embedded in the tip of the inserted needle to assess accuracy RMS after completing the insertion. We scanned the model with CT at the end-inspiratory phase in the university affiliated Renji Hospital. The image resolution is 512 × 512 × 285, and the spacing is 0.7 mm × 0.7 mm × 1.3 mm.

Applied to the robotic puncture system, we carried out the puncture experiment on the respiratory motion simulation phantom, as shown in Figure 10.

In the image coordinate and EMT coordinate, we selected the corresponding fiducial landmarks. Subsequently, we controlled the puncture robot to move in the effective range of the EMT, and recorded four different fiducial landmarks around the model by using the puncture needle. Thus, we realized the registration of the image coordinate, the EMT coordinate and the UR coordinate.

According to the target in the image, we selected the suitable entry point manually and planned the optimal puncture path, as shown in Figure 11a. And the distance between the entry point and the target was 61.5 mm. The robot dynamically adjusted the pose of the needle, and displayed the distance from the needle tip to the target in real time (as shown in the green area in Figure 11b).

Running the respiration motion phantom to simulate the thoracic-abdominal movement, we adopted the principal component of time-series features of external surrogate signals to predict the trajectory of the internal tumour target, and restored the principal component of predicted value of the target to 3D space position. Considering that it took time for a robot to complete an instruction, the refreshing rate of the prediction process and the control command circle for the UR3 motion control were 1 s. Subsequently, the puncture needle dynamically tracked the target's movement in the respiratory movement phantom. The displacement amplitude of the abdominal marker and target point tracked by EMT were 1.82 and 2.66 cm, respectively. According to optimal puncture planning path, the puncture robot can finally automatically insert the needle into the target at the same timing point, as shown in Figure 12.

Using the same path, the experiment was repeated five times, where the average RMS, TRE of $T_{Mag \leftarrow CT}$, $T_{UR \leftarrow Mag}$ were shown in Table 3, and the RMS of the puncturing error was 0.65 mm.

4 | DISCUSSION

By experimenting with both the public datasets and respiratory motion simulation phantom, we verify the improved property of the proposed framework for target localization based on LSTM. Compared with the methods of establishing a correlation model or a prediction model based on the external surrogate signals and the trajectory of the internal tumour target, respectively, our method comprehensively integrates the correlation model and prediction model together, which not only reduces the number of models but also improves tracking accuracy.

The public dataset study verified the performance of the proposed prediction method. Assume that the test set should derive from the same distribution as training data, they are divided according to each patient. And we can determine the value of key parameters in the framework for respiration prediction. The optimal size of the sliding window which represents the number of historical data to make prediction is 30, and the optimal number of LSTM layers is 1. We compared LSTM with some other RNN networks including Bi-LSTM, GRU and Bi-GRU in the prediction performance under the same super parameters. The experimental results reveal

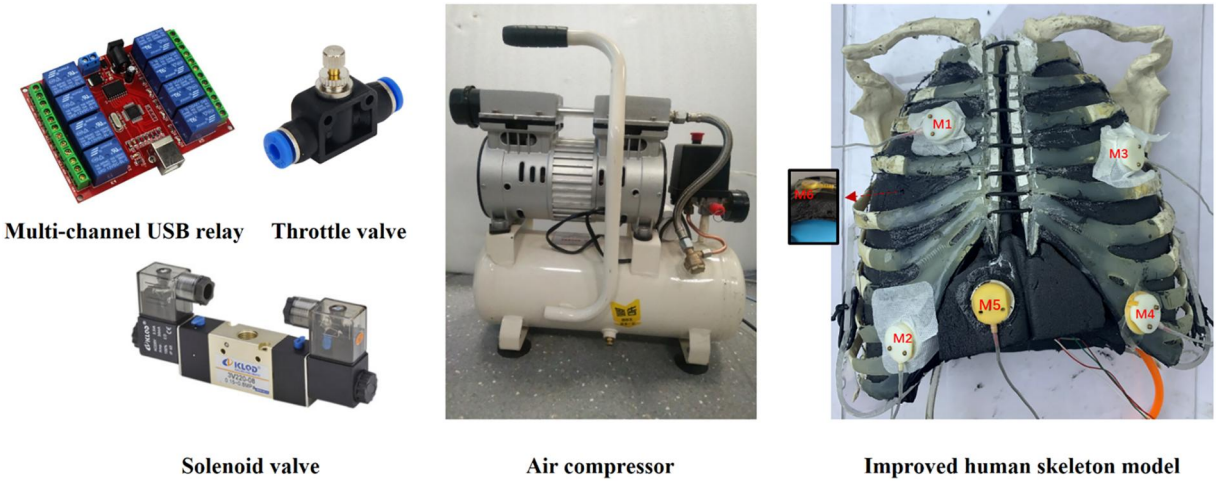

FIGURE 9 The respiratory motion simulation phantom

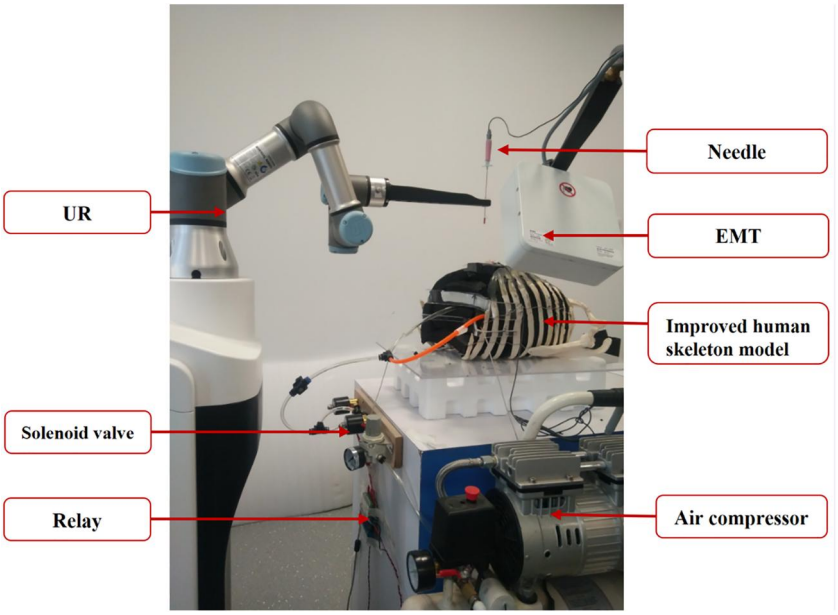

FIGURE 10 The set-up of the respiration motion phantom experiment

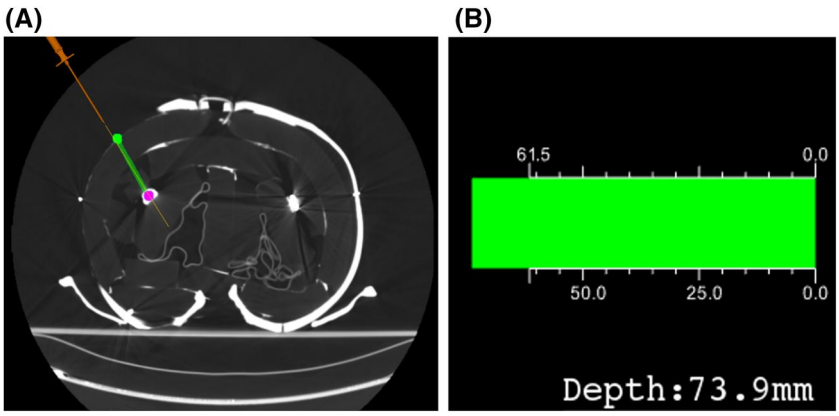

FIGURE 11 (A) The planned optimal puncture path was displayed. The green point and pink point represent the selected entry point and the target, respectively. (B) Distance from the needle tip to the target was measured

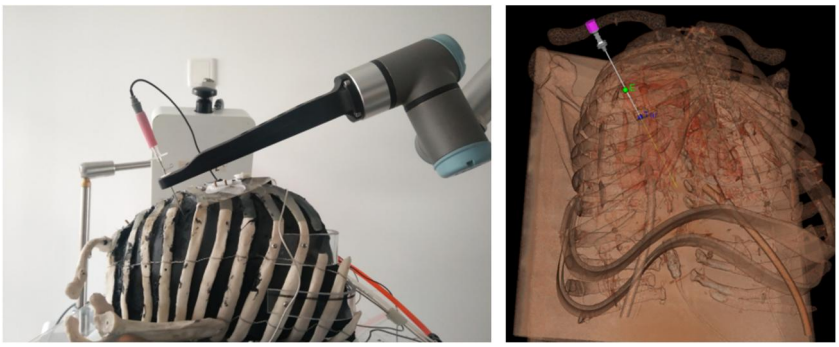

FIGURE 12 The result of target puncturing. The green point and blue point represent the selected entry point and the target at which the needle reaches, respectively

TABLE 3 Average errors of target localization in the respiratory motion phantom

Number of times	$T_{Mag \leftarrow CT}$ RMS (mm)	$T_{Mag \leftarrow CT}$ TRE (mm)	$T_{UR \leftarrow Mag}$ RMS (mm)	$T_{UR \leftarrow Mag}$ TRE (mm)	Puncturing RMS (mm)
1	0.73	1.21	0.21	0.32	0.56
2	0.91	1.54	0.42	0.55	0.72
3	0.88	1.46	0.34	0.48	0.64
4	0.79	1.39	0.22	0.41	0.59
5	0.98	1.68	0.48	0.59	0.75
Average	0.86	1.46	0.33	0.47	0.65

Abbreviation: RMS, root mean square.

that LSTM achieved a higher precision, with an MAE of 0.44 mm and an RMSE of 0.58 mm. Thus, we adopt LSTM as the framework for respiratory prediction. Besides, compared with an existing method ε-SVR, our method is superior to it with the better characteristics of fully extracting time-series data.

In respiratory motion simulation phantom study, the robotic puncture system can automatically adjust the needle insertion pose, target the tumour's movement and insert the needle into the target according to the predicted value. The average RMS of the puncturing error is 0.65 mm. The accuracy of multi-coordinate registration has a positive effect on target localization for reducing puncturing errors. During the procedure, the needle can bend when it is inserted into the moving phantom. Because a sensor was embedded in the tip of the needle, the localization accuracy is not influenced. The robotic puncture system helps surgeon to be less dependent on experience, when performing 3D reconstruction of CT scan images for surgical planning, the angular adjustment of puncture needle, and localization of the drift target. With its stable mechanical structure, the robotic puncture system avoids most human factors, such as trembling of hands, and improves one-shot success rate.

Since percutaneous surgery is carried out under the precondition of regular breath which represent most clinical cases, we utilize the relatively stable part of the public datasets and adopt regular pattern to simulate respiratory system by air exchange in the moving phantom. Due to the budget limitation, we failed to perform prediction on more complex respiration motions, such as from regular breathing to irregular breathing caused by coughing and sneezing. In addition, it is also necessary to consider the differences of respiratory movement caused by people from different age or sex groups and operating positions. Besides, the safety protection of the puncture robot should be taken into account. We realize that the robotic arm can encounter the phantom sometimes and the series manipulator has a large angle deflection in the minor range of motion, which is likely to cause damage to the phantom. Thus, it is necessary to monitor the movement path of UR3 and avoid the risk of operation. When the system function is improved, the accuracy of the method will be further verified through animal experiments.

5 | CONCLUSION

Focusing on the drift in targeting position during the procedure, we proposed a framework for target localization based on LSTM. We combine the correlation model and the prediction model by using LSTM, and adopt the principal component of time-series features of external surrogate signals to predict the trajectory of the internal tumour target, which is validated in public datasets. Additionally, based on the UR3 robot arm and EMT, we applied the proposed approach to a prototype of robotic puncture system for real-time free-breath tumour tracking. The system is tested on customized motion phantom, which achieved a better accuracy of target localization and appealed the potentials applying to clinical application. As the future work, we will continue to increase the accuracy and the robustness of the proposed method on more

clinical data and more complex respiration motions. Besides, it is necessary to improve the functions of the robotic puncture system, especially the safety protection, to carry out animal experiments.

ACKNOWLEDGEMENT

This research is partially supported by the National Key research and development program (2016YFC0106200), Beijing Municipal Natural Science Foundation (L192006), and the funding from IMR of SJTU as well as the 863 national research fund (2015AA043203).

CONFLICT OF INTERESTS

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Institute of Robotics and Cognitive Systems, Lubeck University, Germany, at http://dx.doi.org/10.1109/EMBC.2013.6610839.

ORCID

Yuxiang Ma https://orcid.org/0000-0001-9505-4594
Lixu Gu https://orcid.org/0000-0002-6210-4847

REFERENCES

1. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med*. 2011;32(4):605-644.
2. Xing L, Thorndyke B, Schreibmann E, et al. Overview of image-guided radiation therapy. *Med Dosim*. 2006;31(2):91-112.
3. Lee S, Yang DS, Choi MS, Kim CY. Development of respiratory motion reduction device system (RMRDs) for radiotherapy in moving tumors. *Jpn J Clin Oncol*. 2004;34(11):686-691.
4. Keall P, Yamamoto T, Suh Y. Introduction to 4D motion modeling and 4D radiotherapy. In: Ehrhardt J, Lorenz C, eds. *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*. Springer Berlin Heidelberg; 2013:1-21.
5. Kilby W, Dooley JR, Kuduvalli G, Sayeh S, Maurer CR , Jr. The CyberKnife robotic radiosurgery system in 2010. *Technol Cancer Res Treat*. 2010;9(5):433-452.
6. Kilby W, Naylor M, Dooley JR, Maurer CR, Sayeh S. 2 - A technical overview of the CyberKnife system. In: Abedin-Nasab MH, ed. *Handbook of Robotic and Image-Guided Surgery*. Elsevier; 2020:15-38.
7. Ernst F, Martens V, Schlichting S, et al. Correlating chest surface motion to motion of the liver using ε-SVR–a porcine study. In: *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*; 2009:356-364.
8. Ernst F, Bruder R, Schlaefer A, Schweikard A. Correlation between external and internal respiratory motion: a validation study. *Int J Comput Assist Radiol Surg*. 2012;7(3):483-492.
9. Seregni M, Pella A, Riboldi M, Orecchia R, Cerveri P, Baroni G. Real-time tumor tracking with an artificial neural networks-based method: a feasibility study. *Phys Med*. 2013;29(1):48-59.
10. Dou M, Yu S, Sun R, Wang C, Sun L. UT transform based tumor respiratory motion estimation and prediction for radiosurgery robot. In: *Proceedings of 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*; 2017:1081-1086.
11. Hong SM, Jung BH, Ruan D. Real-time prediction of respiratory motion based on a local dynamic model in an augmented space. *Phys Med Biol*. 2011;56(6):1775-1789.
12. Sun WZ, Jiang MY, Ren L, Dang J, You T, Yin FF. Respiratory signal prediction based on adaptive boosting and multi-layer perceptron neural network. *Phys Med Biol*. 2017;62(17):6822-6835.
13. Wang R, Liang X, Zhu X, Xie Y. A feasibility of respiration prediction based on deep Bi-LSTM for real-time tumor tracking. *IEEE Access*. 2018;6:51262-51268.
14. Graves A, Mohamed A-r, Hinton G. Speech recognition with deep recurrent neural networks. In: *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*; 2013:6645-6649.
15. Li K, Daniels J, Liu C, Herrero P, Georgiou P. Convolutional recurrent neural networks for glucose prediction. *IEEE J-BHI*. 2020;24(2):603-613.
16. Kim B, Kang CM, Kim J, Lee SH, Chung CC, Choi JW. Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network. In: *Proceedings of 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*; 2017:399-404.
17. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process*. 1997;45(11):2673-2681.
18. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735-1780.
19. Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. *Neural Comput*. 2000;12(10):2451-2471.
20. Wold S, Esbensen K, Geladi PJC. Principal component analysis. *Chemom Intell Lab Syst*. 1987;2(1-3):37-52.
21. https://www.ndigital.com/medical/products/aurora/
22. https://www.universal-robots.com/products/ur3-robot/
23. Berthold KPH. Closed-form solution of absolute orientation using unit quaternions. *J Opt Soc Am*. 1987;4(4):629-642.
24. Fitzpatrick JM, West JB, Maurer CR , Jr. Predicting error in rigid-body point-based registration. *IEEE Trans Med Imaging*. 1998;17(5):694-702.
25. Durichen R, Davenport L, Bruder R, Wissel T, Schweikard A, Ernst F. Evaluation of the potential of multi-modal sensors for respiratory motion prediction and correlation. In: *Proceedings of 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; 2013:5678-5681.
26. Ernst F. *Compensating for Quasi-Periodic Motion in Robotic Radiosurgery*. Springer Science & Business Media; 2011.
27. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36(8):1627-1639.
28. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw* 2005;18(5-6):602-610.
29. Chung J, Gulcehre C, Cho K, Yoshua B. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014. arXiv:1412.3555.
30. Zhao R, Wang D, Yan R, Mao K, Shen F, Wang J. Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Trans Ind Electron*. 2018;65(2):1539-1548.

How to cite this article: Ma Y, Yang Z, Wu W, Xie H, Gu L. Target localization during respiration motion based on LSTM: A pilot study on robotic puncture system. *Int J Med Robot*. 2021;17:e2247. https://doi.org/10.1002/rcs.2247