


RESEARCH

Open Access



# Real-time liver tracking algorithm based on LSTM and SVR networks for use in surface-guided radiation therapy

Guangyu Wang<sup>1</sup>, Zhibin Li<sup>1</sup>, Guangjun Li<sup>1\*</sup> , Guyu Dai<sup>1</sup>, Qing Xiao<sup>1</sup>, Long Bai<sup>1</sup>, Yisong He<sup>1</sup>, Yaxin Liu<sup>1,2</sup> and Sen Bai<sup>1</sup>

## Abstract

**Background:** Surface-guided radiation therapy can be used to continuously monitor a patient's surface motions during radiotherapy by a non-irradiating, noninvasive optical surface imaging technique. In this study, machine learning methods were applied to predict external respiratory motion signals and predict internal liver motion in this therapeutic context.

**Methods:** Seven groups of interrelated external/internal respiratory liver motion samples lasting from 5 to 6 min collected simultaneously were used as a dataset,  $D_v$ . Long short-term memory (LSTM) and support vector regression (SVR) networks were then used to establish external respiratory signal prediction models (LSTMpred/SVRpred) and external/internal respiratory motion correlation models (LSTMcorr/SVRcorr). These external prediction and external/internal correlation models were then combined into an integrated model. Finally, the LSTMcorr model was used to perform five groups of model updating experiments to confirm the necessity of continuously updating the external/internal correlation model. The root-mean-square error (RMSE), mean absolute error (MAE), and maximum absolute error (MAX\_AE) were used to evaluate the performance of each model.

**Results:** The models established using the LSTM neural network performed better than those established using the SVR network in the tasks of predicting external respiratory signals for latency-compensation (RMSE < 0.5 mm at a latency of 450 ms) and predicting internal liver motion using external signals (RMSE < 0.6 mm). The prediction errors of the integrated model (RMSE  $\leq$  1.0 mm) were slightly higher than those of the external prediction and external/internal correlation models. The RMSE/MAE of the fifth model update was approximately ten times smaller than that of the first model update.

**Conclusions:** The LSTM networks outperform SVR networks at predicting external respiratory signals and internal liver motion because of LSTM's strong ability to deal with time-dependencies. The LSTM-based integrated model performs well at predicting liver motion from external respiratory signals with system latencies of up to 450 ms. It is necessary to update the external/internal correlation model continuously.

**Keywords:** Radiotherapy, Respiratory motion, Liver tracking, LSTM, SVR, Prediction

## Introduction

During the radiotherapy of thoracic–abdominal tumors, respiratory motion can cause tumor displacement that affects the accuracy of radiotherapy [1]. In particular, stereotactic body radiation therapy (SBRT) technique [2] has increased requirements in terms of irradiation accuracy

\*Correspondence: gjnick829@sina.com

<sup>1</sup> Department of Radiation Oncology, Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China

Full list of author information is available at the end of the article



© The Author(s) 2020. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

and, therefore, will be impacted more by respiratory motion. To manage the tumor motion caused by respiration, motion-encompassing [3, 4] and breath-holding methods [5–8] as well as forced shallow-respiratory with abdominal compression [9], respiratory-gating [10, 11], and respiration-synchronized techniques [12] are used clinically. Both respiratory-gating and respiration-synchronized techniques require the real-time tracking of tumors. Tumor tracking techniques can be divided into two categories. The first is direct real-time tracking technique, in which X-ray imaging is used to locate tumors or implanted metal markers, or electromagnetic methods are used to track implanted coils in the target volume. These methods either require additional irradiation or are invasive [13–16]. The second approach involves the indirect real-time tracking technique, in which tumor position is predicted based on external surrogate respiratory signals acquired using optical or infrared devices [17, 18] or spirometry [19]. Owing to its non-radiological and non-invasive nature, the indirect real-time tracking technique can be applied in the clinic using surface-guided radiation therapy (SGRT). This technique has many advantages such as imaging without the requirement of dosing, real-time feedback, 3D sub-millimeter spatial resolution, non-invasive and non-contact application, ease of use, and an enhanced field of view (FOV) [20]. These advantages make SGRT well-suited to real-time tracking and respiratory gating, and many studies have reported applications of SGRT in the treatment of lung [21] and breast tumors [22], laryngeal cancer [23], etc.

Establishing an accurate external respiratory signal prediction model and an appropriate external/internal correlation model between external surrogates and internal tumor motion is the key to the successful implementation of indirect real-time tumor tracking technology [24]. The external respiratory signal prediction and external/internal respiratory motion correlation models allow for the compensation of system latency and the prediction of internal tumor motion, respectively. In respiratory gating in clinical settings, the latency of SGRT can result in treatment inefficiencies and/or geographic misses [25]. According to the AAPM TG 142 report [26], latency should be within 100 ms, a tolerance that assumes a moving object traveling at a speed that is not higher than 20 mm/s, corresponding to a positional uncertainty of 2 mm [26]. However, the total latency measured in clinical practice is often greater than 100 ms, and cases involving latencies of 200–400 ms are often reported [27, 28]. In such cases, the positional uncertainty of the tumor can reach more than 8 mm, which can cause unexpected loss of target dose coverage and result in cold spots [29]. Moreover, the positional uncertainty caused

by high latency can also lead to interplay effects, blurring and spatial deformation of the dose distribution [30], and substantial discrepancies between planned and delivered doses [31]. The system latency is inherent in a real-time tracking system associated with the processes of obtaining external imaging data, predicting internal tumor position from the external/internal correlation model, and adjusting the radiation beam accordingly [32]. Different studies have shown that the system latencies of different real-time tracking devices range from tens of milliseconds to more than 400 ms [32–34]. As mentioned previously, a latency of 100 ms will result in a tumor positional uncertainty of 2 mm [26].

External respiratory motion prediction algorithms developed in previous studies can be divided into two categories. The first include algorithms based on existing models, primarily the least-squares fitting algorithm [35], algorithms based on the extended Kalman filter [36], and fuzzy logic algorithms [17, 37]. Such algorithms are based on the assumption that the existing respiratory pattern remains stationary and periodic—an assumption that might be incorrect [24]. The second category concerns model-free prediction algorithms, including neural networks [38–40], regression analysis [41], and support vector regression (SVR) [42]. The advantage of these algorithms is that they do not require a stable respiratory pattern.

External/internal correlation algorithms developed in other studies primarily apply linear [43, 44], piecewise linear [45], polynomial [46], and combined polynomial [47] models or adaptive filters, neural networks [48, 49], or SVR [50]. However, relatively simple linear polynomial models suffer from poor robustness and produce relatively large prediction errors for samples with large volatility [46]. Algorithms such as filters, neural networks such as long short-term memory (LSTM), and SVR networks are stronger at processing nonlinear problems and have better real-time dynamic prediction performance. The LSTM neural network was first proposed by Hochreiter and Schmidhuber [51] in 1997. It has been recognized for its outstanding performance in processing sequence information and is widely used in speech recognition and machine translation tasks [52, 53]. SVR is an important application branch of support vector machines (SVMs) that is used to minimize structural risks and seek globally optimal solutions. SVR applies nonlinear transformation to convert actual nonlinear problems into high-dimensional feature spaces in which the nonlinear problems can be solved by constructing linear decision functions [54]. Based on the characteristics and advantages of LSTM and SVR mentioned above, the LSTM and SVR networks were used to establish an external/internal respiratory motion correlation model (external/

internal correlation model) and an external respiratory signals prediction model (external prediction model) in this study.

Indirect real-time tracking techniques based on external surrogate respiratory signals are completely free of additional radiation and are noninvasive. To implement these techniques, an external prediction model must be combined with an external/internal correlation model. All of the previous studies in this area were limited to external respiratory motion or internal tumor/organ motion prediction alone and did not reflect actual clinical situations, in which latency compensation and the prediction of internal tumor position should be performed simultaneously. Considering the limitation of the previous studies above, these models were combined into an integrated model for clinical practice in this study. The integrated model uses external respiratory motion signals to predict internal liver motion and compensates for system latency. The performance of the respective networks and their combination was compared and analyzed. Finally, to reflect the temporally changing relationship between external and internal respiratory motion [55, 56], the necessity of continuously updating the model was also verified.

## Methods

### Dataset

The respiratory motion data of seven volunteers collected by the Institute of Robotics and Cognitive Systems of the University of Lübeck in Germany [47, 57] were used. The respiratory motion signals of the chest surface and internal liver were collected in the x- (left–right), y- (superior–inferior), and z- (anterior–posterior) directions, respectively. The samples collected for each volunteer were taken as individual sets of data, which were combined to form the final dataset,  $D_v$ . The respiratory motion signals of the internal liver were obtained using 4D ultrasound and template matching to track the motion of the liver vessel bifurcation point. In the template matching process, phase-only correlation (POC) [58] was performed followed by normalized cross-correlation (NCC) [59] with interpolation on the target volume to determine the point of optimal registration; details on this process can be found in [47]. External respiratory motion samples were obtained using an AccuTrack 250 system to track the motion of LEDs fixed on the surface of each volunteer's chest [47]. The duration of respiratory motion sample collection for each set was 5–6 min, and approximately 6500 to 8100 sampling points were obtained in each direction per set. The sampling frequency of the internal respiratory motion sample was 20 Hz. Each external respiratory motion sample was resampled to make it consistent in length with the

corresponding internal respiratory motion sample. Prior to model training, preprocessing operations such as outlier deletion, smoothing, filtering, and standardization were performed on  $D_v$ .

### LSTM network and SVR network

#### LSTM network

Sets of training and training label samples ( $X_{\text{train}} = [x_1, x_2, \dots, x_m]$  and  $Y_{\text{train}} = [y_1, y_2, \dots, y_m]$ , respectively) and testing and testing label samples ( $X_{\text{test}} = [x_{m+1}, x_{m+2}, \dots, x_{m+n}]$  and  $Y_{\text{test}} = [y_{m+1}, y_{m+2}, \dots, y_{m+n}]$ , respectively) were obtained for the LSTM network. The LSTM network comprises multiple memory blocks, each of which contains one cell and three gates (input, forget, and output gates) [60]. Through a forward propagation process, the LSTM network turns input data into output samples. The entire forward propagation of one block from input to output is carried out as follows:

The output  $b_i^t$  of the input gate is

$$b_i^t = f(a_i^t), \quad (1)$$

where

$$a_i^t = \sum_{i=1}^I w_{il}x_i^t + \sum_{c=1}^C w_{cl}s_c^{t-1} + \sum_{h=1}^H w_{hl}b_h^{t-1}. \quad (2)$$

The output  $b_\phi^t$  of the forget gate is

$$b_\phi^t = f(a_\phi^t), \quad (3)$$

where

$$a_\phi^t = \sum_{i=1}^I w_{i\phi}x_i^t + \sum_{c=1}^C w_{c\phi}s_c^{t-1} + \sum_{h=1}^H w_{h\phi}b_h^{t-1}. \quad (4)$$

The output  $b_w^t$  of the output gate is

$$b_w^t = f(a_w^t), \quad (5)$$

where

$$a_w^t = \sum_{i=1}^I w_{iw}x_i^t + \sum_{c=1}^C w_{cw}s_c^t + \sum_{h=1}^H w_{hw}b_h^{t-1}. \quad (6)$$

From these outputs, the output  $s_c^t$  of the cell is obtained as

$$s_c^t = b_\phi^t s_c^{t-1} + b_i^t g(a_c^t), \quad (7)$$

where

$$a_c^t = \sum_{i=1}^I w_{ic}x_i^t + \sum_{h=1}^H w_{hc}b_h^{t-1}. \quad (8)$$

the final output of the block,  $b_c^t$ , is then calculated as the product of  $b_w^t$  and a function of  $s_c^t$ :

$$b_c^t = b_w^t h(s_c^t). \tag{9}$$

It should be noted that, in the above equations,  $f$ ,  $g$ , and  $h$  denote the activation functions used by the three gates, the input, and the final output, respectively.

**SVR network**

Sets of training and testing samples ( $D_{\text{train}} = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m), y_i \in \mathbb{R}\}$  and  $D_{\text{test}} = \{(x_{m+1}, y_{m+1}), (x_{m+2}, y_{m+2}), \dots, (x_{m+n}, y_{m+n}), y_i \in \mathbb{R}\}$ , respectively) were obtained for the SVR network, which attempts to find a model  $f(x)$  in which  $y^* = f(x_i)$  and  $y_i$  are as close as possible [54, 61]. For a maximum tolerable deviation between  $y^*$  and  $y_i$  of  $\epsilon$ , the SVR problem can be formalized as

$$\min_{w,b} \frac{1}{2} \hat{\alpha} \hat{L} - w \hat{\alpha} \hat{L}^{-2} + C \sum_{i=1}^m l_{\epsilon}(f(x_i), y_i), \tag{10}$$

where  $C$  is the regularization constant and  $l_{\epsilon}$  is the  $\epsilon$ -insensitive loss function.

By introducing the slack variables  $\xi_i$  and  $\hat{\xi}_i$ , Eq. (10) can be expressed as

$$\min_{w,b,\xi_i,\hat{\xi}_i} \frac{1}{2} \hat{\alpha} \hat{L} - w \hat{\alpha} \hat{L}^{-2} + C \sum_{i=1}^m (\xi_i, \hat{\xi}_i). \tag{11}$$

The Lagrange multiplier  $u_i$  can then be introduced to obtain the SVR solution as

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x + b, \tag{12}$$

where  $b$  is the model parameter to be determined and  $f(x)$  is the final model found by the SVR method.

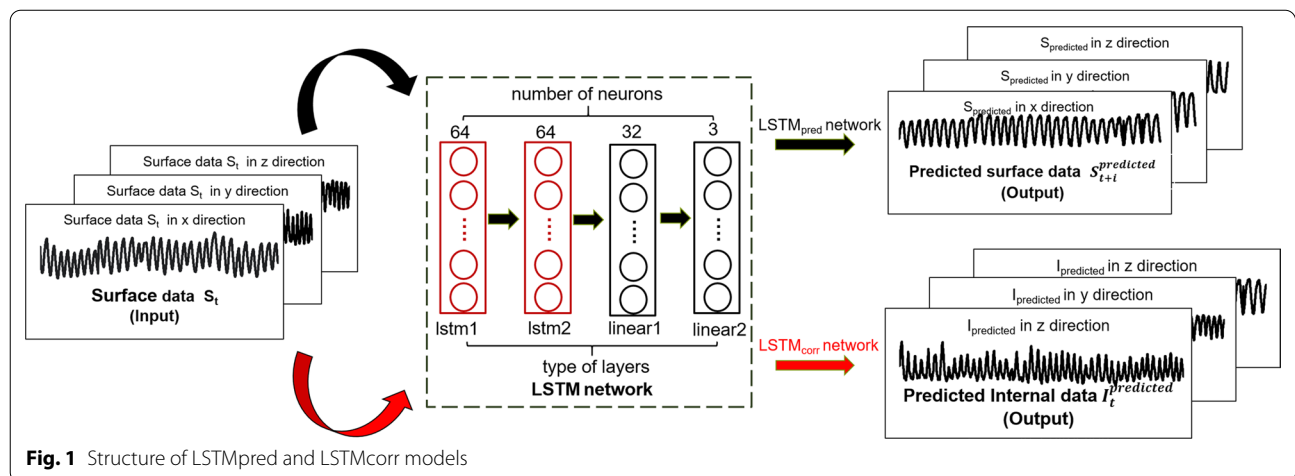
**External respiratory signals prediction models**

The LSTM prediction (LSTMpred) and SVR prediction (SVRpred) models were established using the LSTM neural and SVR networks, respectively, to predict the respiratory motion of the chest surface at the system latency. Both LSTMpred and SVRpred use the current external respiratory motion signals  $S_t$  to predict the future external respiratory motion signals  $S_{t+i}$ . Considering that the length of the training and label datasets should be the same during the training of the SVR model, a  $S_t$ -to- $S_{t+i}$  length ratio of 1:1 was selected for both models. As the typical system latency ranges from tens of milliseconds to more than 400 ms, the respiratory motion prediction algorithm results were assessed at latencies of 50, 150, 200, and 450 ms, i.e., the  $i=1, 3, 4,$  and  $9$  latencies were set to 50, 150, 200, and 450 ms, respectively.

The external respiratory motion data of dataset  $D_v$  were divided at ratios of 9:1 into training and testing data for the LSTMpred and SVRpred models. After testing, comparison, and adjustment to determine the best network performance, appropriate parameters were selected for the two external prediction models. For the LSTMpred model, mean square error (MSE) loss was chosen as the loss function and the network was trained using an Adam optimizer with a learning rate of 0.001 and a batch size of 64. A dropout rate of 20% over 60 total epochs and 20 time steps was used in the LSTMpred model adjustment process. The number of layers and neurons in each layer are shown in Fig. 1. For the SVRpred model, the radial basis function (RBF) kernel was selected as the kernel function, and the gamma and penalty parameter  $C$  of SVRpred were set as 0.1 and 1000, respectively.

**External/internal respiratory motion correlation models**

The LSTM and SVR networks were used to establish the external/internal correlation models LSTMcorr and



**Fig. 1** Structure of LSTMpred and LSTMcorr models

SVRcorr, respectively, which use the current external respiratory motion signal  $S_t$  to predict the current internal liver respiratory motion sample  $I_t$ . The preprocessed dataset  $D_v$  was divided at 9:1 ratios into training and testing data for the LSTMcorr and SVRcorr models. Because the external and internal samples had the same length and corresponded to each other, a  $S_t$ -to- $I_t$  ratio of 1:1 was selected for both models.

The structure and parameter selection processes applied in obtaining the external/internal correlation models were equivalent to those used for the external prediction models except that 40 instead of 60 epochs were used to train LSTMcorr. The external/internal correlation and external prediction models differ primarily in terms of the training and label data used to train the respective models. Both sets of models were developed based on a Pytorch deep learning framework using Python.

### Integrated model

The integrated model derived as a combination of the trained external prediction and trained external/internal correlation models uses external respiratory motion signals to predict internal liver motion while compensating for system latency. The efficiency of this combination of the two models was assessed by predicting future internal liver positions.

### Verification of external/internal correlation model update

During free breathing, different parts of the anatomy can move with different temporal and spatial relationships that will change continuously over time [24, 62]. Unexpected behaviors such as talking and coughing can also induce obvious changes in these relationships. Therefore, the established external/internal correlation model should also change over time. Obtaining the latest sample and updating the model continuously is a common approach to solving this problem [24]. In general, the changes in the temporal and spatial relationships will be slow, and approximately 30 s will be required to detect obvious alterations [48]. Previous studies have shown that updating every 10 s is sufficient to ensure the accuracy of a model [43]. In the model update process, the weights of each computing node within the neural network are updated automatically by entering the latest collected data as input to the network, which does not change the structure of the network or the set parameters.

Although the changing characteristics of a patient's respiratory motion pattern over time will also affect the prediction of external respiratory motion, the influence of this is primarily manifested on a time scale that is longer than the system latency. Thus, the update verification

of the model primarily focuses on the external/internal correlation.

To carry out update verification of the LSTMcorr model, the dataset  $D_v$  was divided into six parts—C1, C2, C3, C4, C5, and C6—in the proportion 1:2:2:2:2:1. Five groups of assessments were carried out using C1, C1 + C2, C1 + C2 + C3, C1 + C2 + C3 + C4, and C1 + C2 + C3 + C4 + C5 as the training sets and the unified C6 as the testing set. Each group of assessments was carried out five times. The  $\overline{RMSE}/\overline{MAE}$  of each assessment was taken as the final result for comparison, with the results for each of the five groups normalized to the fifth update to enable an intuitive comparison.

### Evaluation

The root-mean-squared errors (RMSEs), mean absolute errors (MAEs), and maximum absolute errors (MAX\_AEs) of the external prediction, external/internal correlation, and integrated models were used as evaluation indicators to assess the deviations between the predicted and true results,  $\hat{y}_i$  and  $y_i$ , respectively. These indicators are defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (13)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (14)$$

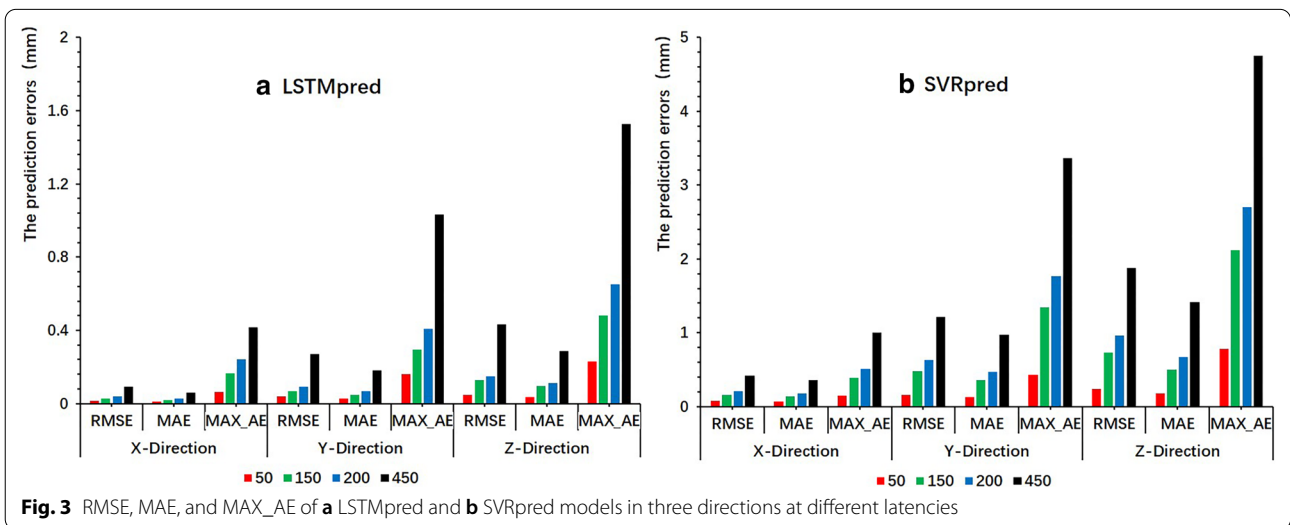
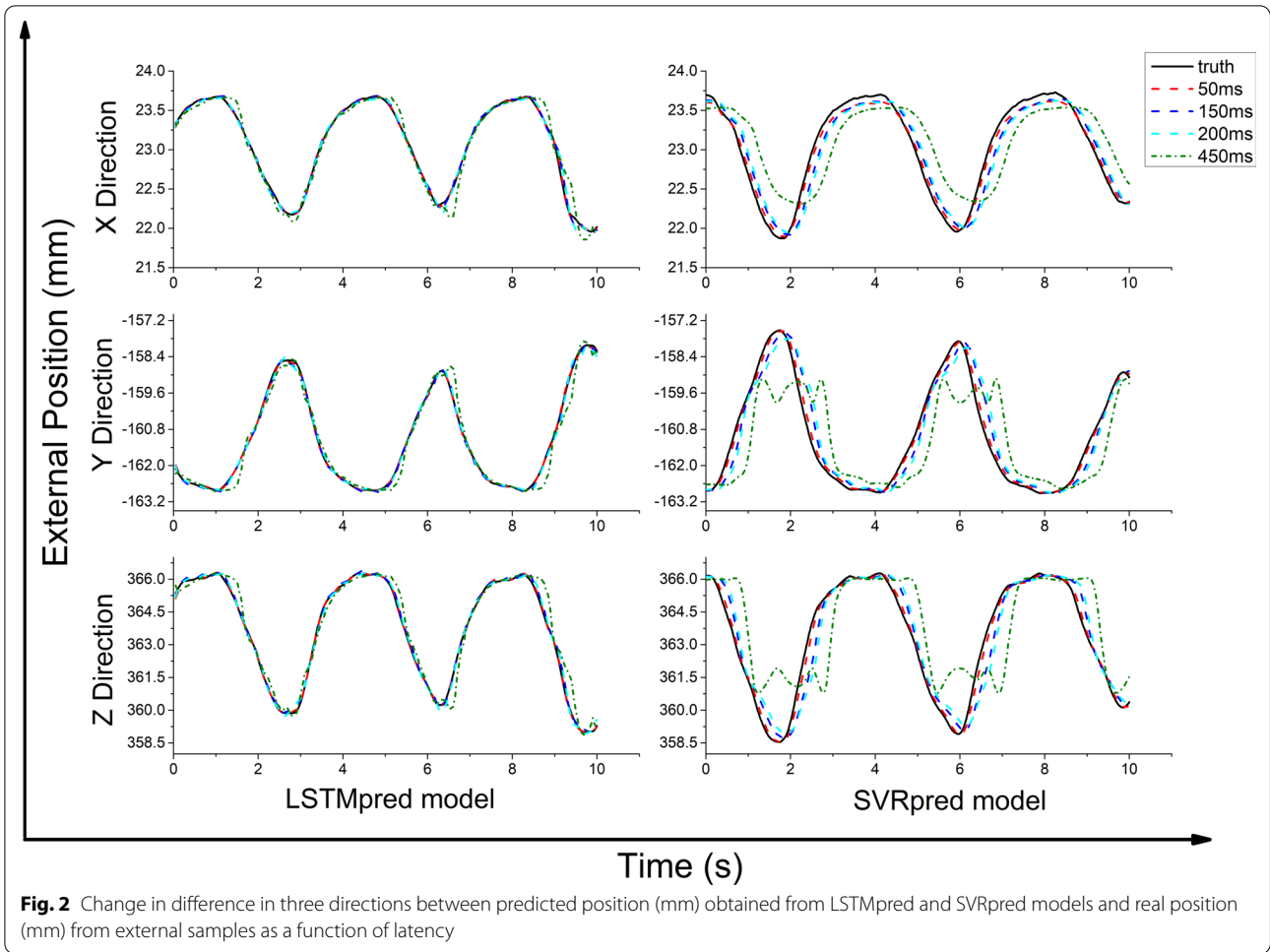
$$MAX\_AE = \max(|y_i - \hat{y}_i|). \quad (15)$$

In each case, a smaller indicator value corresponds to a predicted result that is closer to the real result and a better-performing model.

## Results

### External respiratory motion prediction

The LSTMpred model was found to be much more accurate in all directions than the SVRpred model as an external prediction model (see Figs. 2, 3). For the LSTMpred model with a latency of 450 ms, the maximum prediction errors for all tested cases in the x-, y-, and z-directions were 0.415, 1.034, and 1.529 mm, respectively, while those for the SVRpred model were 1.001, 3.368, and 4.749 mm, respectively. The prediction errors of the LSTMpred and SVRpred models were both largest in the z-direction, followed by the y- and x-directions. The RMSE, MAE, and MAX\_AE values of both models both increased with the latency, indicating that increasing the latency increased the prediction error. This occurred because the human respiratory pattern changes over time



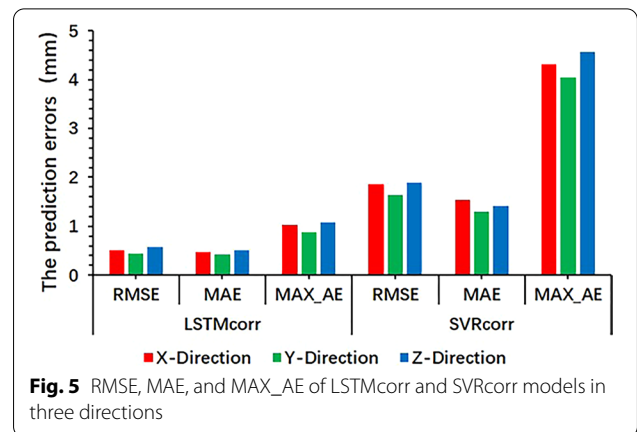
and, therefore, increasing the system latency increases the difference between the predicted and current external respiratory motion signals.

**External/internal correlation model**

The LSTMcorr model was found to be much more accurate in all directions than the SVRcorr model as an external/internal correlation model (see Figs. 4, 5). The RMSE, MAE, and MAX\_AE values of the LSTMcorr model were much smaller than those of the SVRcorr model in all directions, with maximum prediction errors for all tested cases of 1.027, 0.886, and 1.081 mm in the x, y, and z directions, respectively.

**Integrated model**

Because the LSTMpred and LSTMcorr external and external/internal correlation models had smaller errors than the SVRpred and SVRcorr prediction models, the integrated model was constructed using the LSTMpred and LSTMcorr models. The RMSE, MAE, and MAX\_AE of the integrated model were all found to increase with the latency in all directions (see Figs. 6, 7), with RMSE and MAX\_AE values for all tested cases of approximately 1 and 2 mm, respectively. In addition, the prediction errors of the integrated model were slightly greater than those of the external prediction and external/internal

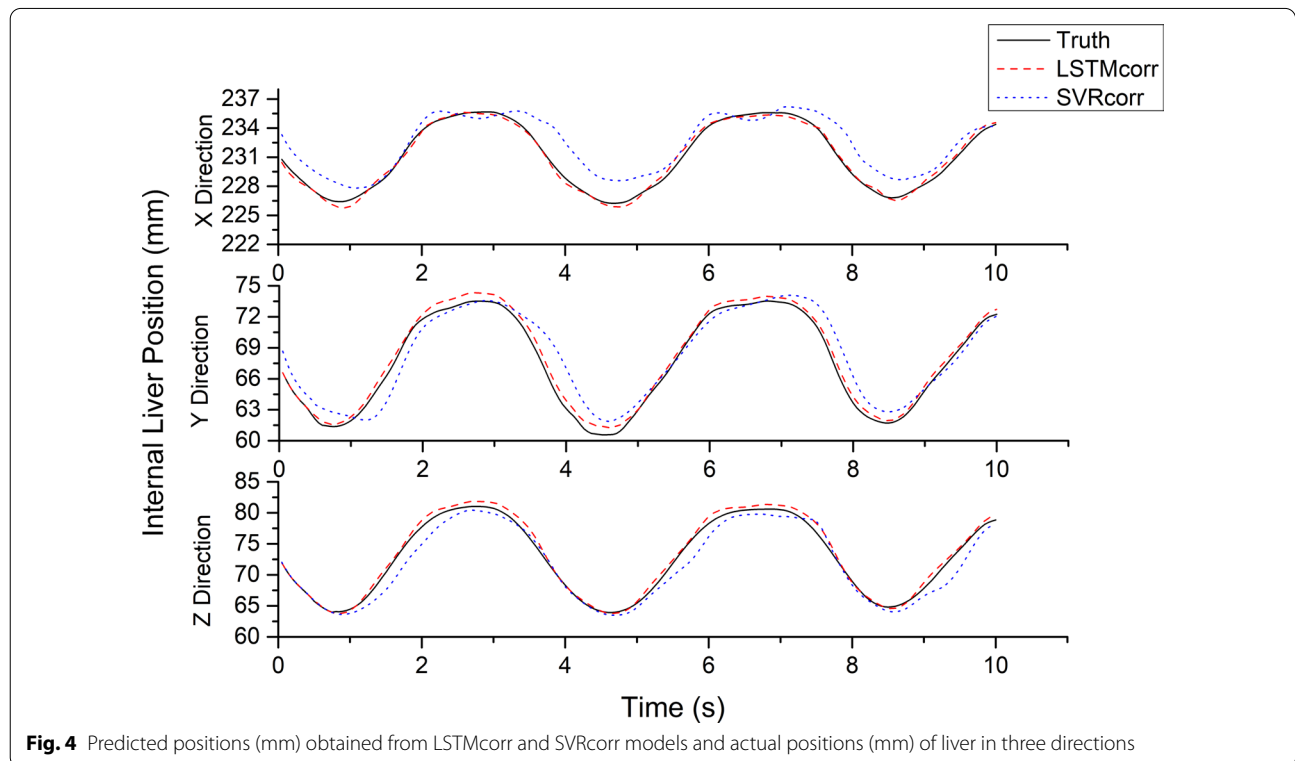


**Fig. 5** RMSE, MAE, and MAX\_AE of LSTMcorr and SVRcorr models in three directions

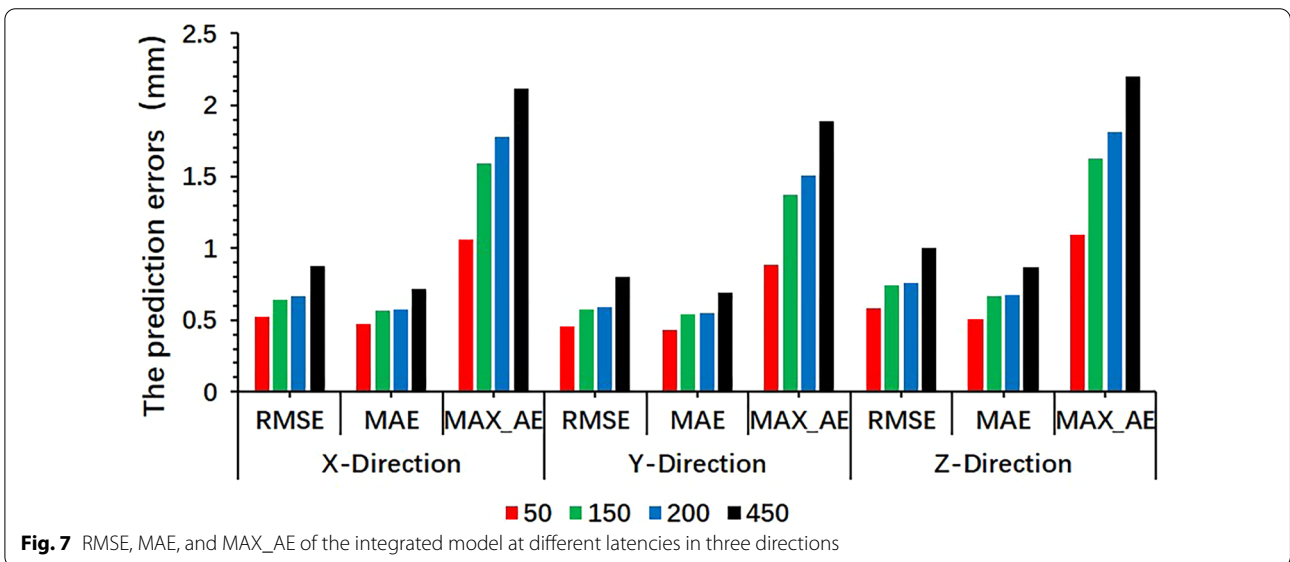
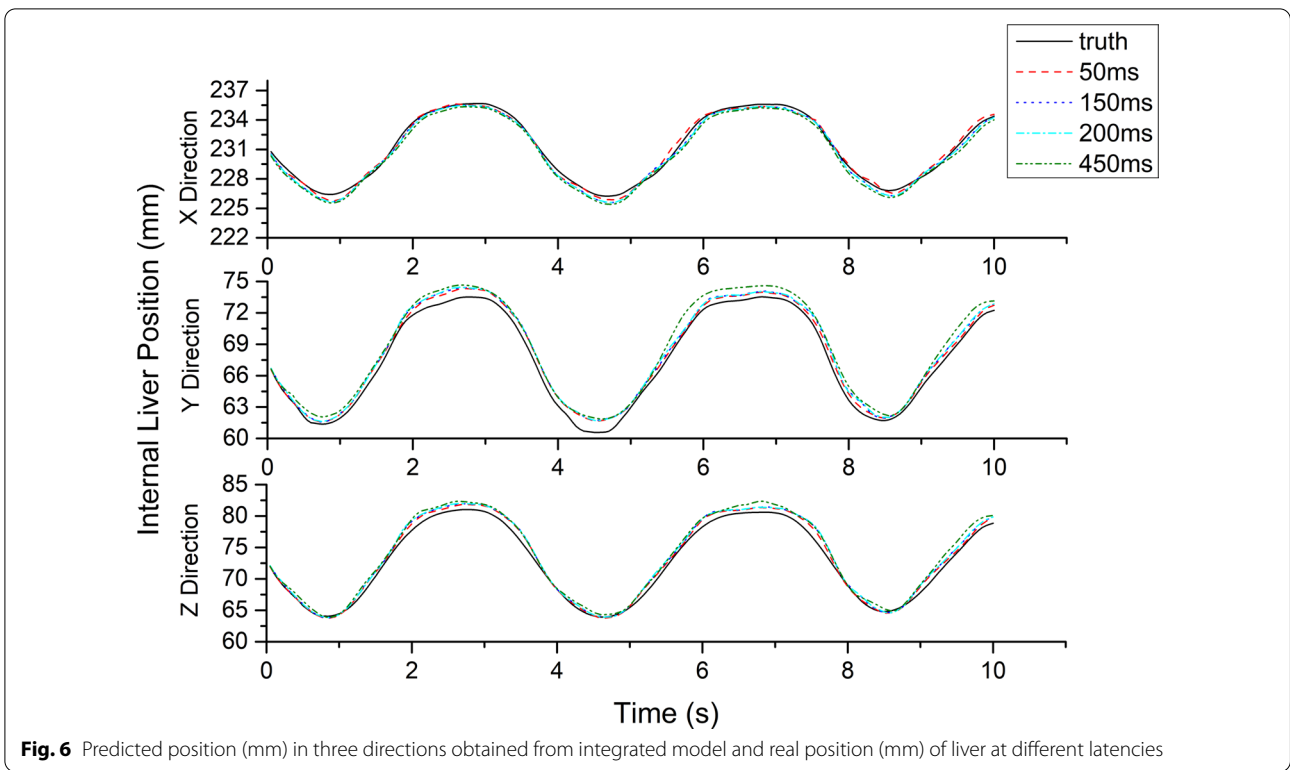
correlation models but smaller than the sum of their errors.

**Verification for external/internal correlation model update**

As the external/internal correlation model is continuously updated, its RMSE and MAE are gradually reduced in all directions (see Fig. 8). As the time interval between the training and test samples decreases, the magnitude of error reduction becomes progressively smaller because reducing the time interval between the two reduces the



**Fig. 4** Predicted positions (mm) obtained from LSTMcorr and SVRcorr models and actual positions (mm) of liver in three directions



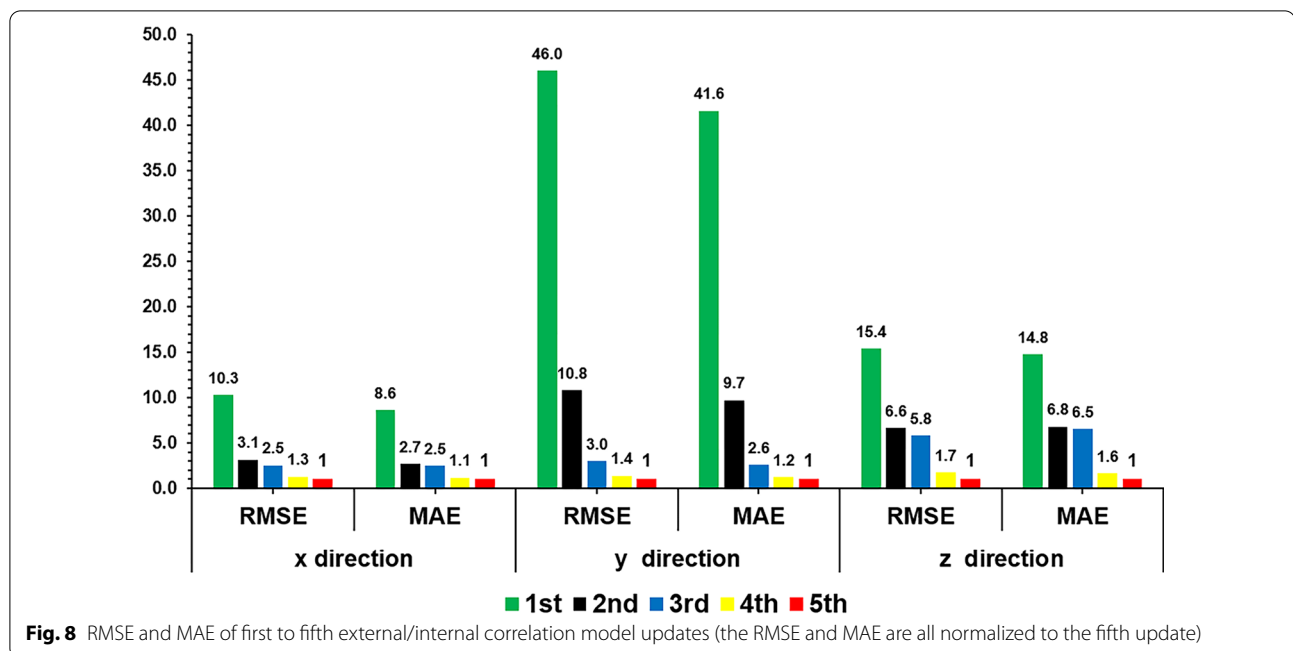
difference in the respiratory motion pattern between the training samples and the test data.

**Discussion**

The external respiratory motion prediction and external/internal correlation models for predicting internal liver motion established by the LSTM neural network

universally outperformed those established by the SVR network in all directions. The primary reason for this is that the LSTM neural network can effectively solve time-dependency problems [63, 64] as a result of the network’s hidden unit, which controls memory and forgetting of time-series information to enable the selective remembering and transmission of these data [65]. In addition,





the nonlinear fitting ability of the SVR network is based on the use of similar history samples, and the SVR cannot reflect further abnormal fluctuations using rare similar history samples [66]. By contrast, LSTM networks have stronger feature extraction capabilities than SVR networks [67–69] and can use these to learn additional features pertaining to abnormal fluctuations. These features provide LSTM networks with more powerful processing capabilities for abnormal fluctuations and gives them smaller prediction errors than SVR networks. Studies in a number of fields, including stock premium [70], snow-melt driven stream flow [71], and traffic flow prediction [69] have shown the superiority of LSTM networks relative to SVR networks in processing time-series data.

In this study, we established an integrated model that predicts internal liver motion from external respiratory signals based on actual clinical tumor real-time motion tracking scenarios and took system latency into account in developing the model. The maximum RMSE and MAX\_AE of the integrated model are approximately 1 and 2 mm, respectively, in all directions, a degree of accuracy that can meet clinical requirements for the real-time motion tracking of liver tumors [72].

A patient's respiratory pattern will change over time as a result of the complex patterns of simultaneous motion of different anatomical structures, which produce phase difference variation and time-changing spatiotemporal correlations [73]. Despite this accepted fact, to date, there have been few relevant external/internal correlation experiments to verify the necessity of model updating in

radiotherapy-based tumor real-time tracking; furthermore, all such studies have been limited to exploring the impact of short-term updates [43, 55]. The results of our model update assessments indicate that the RMSE/MAE obtained for the fifth model update is smaller than that obtained for the first update by an approximate factor of ten in a long-term state. Because the RMSE/MAE decreases as the model is updated, it is necessary to continuously update the model. Doing so ensures that the respiratory motion correlation established by the model matches the situation during the current moment, thereby reducing prediction errors, improving prediction accuracy, and reducing the large impact of random patient actions. Furthermore, the accuracy of the fifth model update is close to that of the fourth update, indicating that when the training/prediction sample length ratio reaches 7:3, a high degree of accuracy can be achieved in the external/internal correlation model.

This study had several limitations. First, the networks employed in the study were specific in that they were all trained based on a respiratory motion dataset obtained from a single institution. Before applying the network developed in this study to any other dataset, therefore, the models' structure and parameters will have to be adjusted and the models will have to be retrained and reevaluated. In addition, the models' performance could be improved by using a larger dataset or a dataset containing multicenter data. Second, the model for predicting internal liver motion from external respiratory signals developed in this study is only suitable when exposed

surface skin immobilization devices such as vacuum bags and stereotactic body frames are used. The model cannot be applied when thermoplastic frame immobilization devices are used because it would not be able to accurately obtain external motion signals and changes in the patient's breathing motion pattern. Third, we only tracked the motion of the liver vessel bifurcation point in our model and, therefore, did not consider nonrigid motion such as deformation of the liver. Finally, a patient's respiratory motion can be severely affected by many factors, including initial nervousness or activity during setup and commencement of treatment followed by eventual relaxation on the table, which alter their breathing pattern. As a result, the application of motion compensation might not always be appropriate, and the question of whether certain patients are suitable for motion compensation treatment alone will have to be explored [74]. In future studies, we will further explore and evaluate the benefits of motion compensation for specific patients.

## Conclusion

In this study, it was shown that LSTM networks outperform SVR networks in the prediction of external respiratory signals and internal liver motion because of the strong ability of the former to capture time dependencies. An LSTM-based integrated model was found to perform well in predicting liver motion from external respiratory signals with a system latency of up to 450 ms. It was also confirmed that continuously updating the model is necessary to improve its prediction accuracy. Our study, could be a meaningful contribution to the real-time motion tracking of liver tumors in clinical practice.

## Abbreviations

FOV: Field of view; LSTM: Long-short term memory; LSTMcorr: The external/internal respiratory motion correlation model established by LSTM network; LSTMpred: The external respiratory signals prediction model established by LSTM network; MAE: Mean absolute errors; MAX\_AE: Maximum absolute errors; MSE: Mean square error; NCC: Normalized cross-correlation; POC: Phase-only correlation; RBF: Radial basis function; RMSE: Root mean squared errors; SBRT: Stereotactic body radiation therapy; SGRT: Surface-guided radiation therapy; SVM: Support vector machine; SVR: Support vector regression; SVRcorr: The external/internal respiratory motion correlation model established by SVR network; SVRpred: The external respiratory signals prediction model established by SVR network.

## Acknowledgements

The authors thank Ernst et al. from University of Lübeck for the respiratory motion data used for this study.

## Authors' contributions

GW and GL conceived and designed the study, drafted the manuscript, and evaluated the result. GL helped given methodological guidance. GW preprocessed experimental data, and established, train, and evaluated the models in this study. ZL helped established and evaluated the models. GW and ZL completed the model updating experiment. GD, QX, and LB analyzed parts of the data, and interpreted the data, and helped evaluated the result. YH, YL, and SB helped evaluated the result. All authors read and approved the final manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81472807 and 81972848).

## Availability of data and materials

The link of datasets used during the current study is available from the paper of references 47.

## Ethics approval and consent to participate

The study was approved by the ethics committee of the West China Hospital.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Department of Radiation Oncology, Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China. <sup>2</sup> College of Physics, Sichuan University, Chengdu 610065, China.

Received: 30 September 2020 Accepted: 6 December 2020

Published online: 14 January 2021

## References

- Tang X, Cullip T, Dooley J, Zagar T, Jones E, Chang S, et al. Dosimetric effect due to the motion during deep inspiration breath hold for left-sided breast cancer radiotherapy. *J Appl Clin Med Phys*. 2015;16:5358.
- McCurdy BMC, McCowan PM. In vivo dosimetry for lung radiotherapy including SBRT. *Phys Med*. 2017;44:123–30.
- Vlachaki M, Castellon I, Leite C, Perkins T, Ahmad S. Impact of respiratory gating using 4-dimensional computed tomography on the dosimetry of tumor and normal tissues in patients with thoracic malignancies. *Am J Clin Oncol*. 2009;32:262–8.
- Colgan R, McClelland J, McQuaid D, Evans PM, Hawkes D, Brock J, et al. Planning lung radiotherapy using 4D CT data and a motion model. *Phys Med Biol*. 2008;53:5815–30.
- Yeoh KW, McNair HA, McDonald F, Hawkins M, Hansen VN, Ramos M, et al. Cone beam CT verification for active breathing control (ABC)-gated radiotherapy for lung cancer. *Acta Oncol*. 2014;53:716–9.
- Wang X. The use of active breathing control (ABC) in radiotherapy of lung cancer. *Chin J Clin Oncol*. 2006;2006:10.
- Wang X, Pan T, Pinnix C, Zhang SX, Salehpour M, Sun TL, et al. Cardiac motion during deep-inspiration breath-hold: implications for breast cancer radiotherapy. *Int J Radiat Oncol Biol Phys*. 2012;82:708–14.
- Bruzzaniti V, Abate A, Pinnaro P, D'Andrea M, Infusino E, Landoni V, et al. Dosimetric and clinical advantages of deep inspiration breath-hold (DIBH) during radiotherapy of breast cancer. *J Exp Clin Cancer Res*. 2013;32:88.
- Media M. A review of kidney motion under free, deep and forced-shallow breathing conditions: implications for stereotactic ablative body radiotherapy treatment. *Technol Cancer Res Treat*. 2014;13:315–23.
- Hu W, Xu A, Li G, Zhang Z, Housley D, Ye J. A real-time respiration position based passive breath gating equipment for gated radiotherapy: a preclinical evaluation. *Med Phys*. 2012;39:1345–50.
- Giraud P, Yorke E, Jiang S, Simon L, Rosenzweig K, Mageras G. Reduction of organ motion effects in IMRT and conformal 3D radiation delivery by using gating and tracking techniques. *Cancer/Radiothérapie*. 2006;10:269–82.
- Buzurovic I, Huang K, Yu Y, Podder TK. A robotic approach to 4D real-time tumor tracking for radiotherapy. *Phys Med Biol*. 2011;56:1299–318.
- Park SH, Kim JC, Kim SJ. Alternative fiducial markers for Vero real-time tumor tracking radiotherapy: a phantom study. *J Korean Phys Soc*. 2016;69:1700–4.
- Shirato H, Harada T, Harabayashi T, Hida K, Endo H, Kitamura K, et al. Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy. *Int J Radiat Oncol Biol Phys*. 2003;56:240–7.

15. Karaman K, Dokdok AM, Karadeniz O, Ceylan C, Engin K. Intravascular placement of metallic coils as lung tumor markers for CyberKnife stereotactic radiation therapy. *Korean J Radiol*. 2015;16:626–31.
16. Oita M, Osaka Y, Onimaru R, Onodera S, Fujita K, Saito E, et al. SU-FF-J-121: retrospective analysis of prostate cancer patients with fiducial gold markers using a real-time tumor tracking system. *Med Phys*. 2006;33:2048.
17. Torshabi AE, Riboldi M, Fooladi AAI, Mosalla SMM, Baroni G. An adaptive fuzzy prediction model for real time tumor tracking in radiotherapy via external surrogates. *J Appl Clin Med Phys*. 2013;14(1):102–14.
18. Torshabi AE, Pella A, Riboldi M, Baroni G. Targeting accuracy in real-time tumor tracking via external surrogates: a comparative study. *Technol Cancer Res Treat*. 2010;9:551–62.
19. Hughes S, McClelland J, Tarte S, Lawrence D, Ahmad S, Hawkes D, et al. Assessment of two novel ventilatory surrogates for use in the delivery of gated/tracked radiotherapy for non-small cell lung cancer. *Radiother Oncol*. 2009;91:336–41.
20. Batista V, Meyer J, Kügele M, Al-Hallaq H. Clinical paradigms and challenges in surface guided radiation therapy: where do we go from here? *Radiother Oncol*. 2020. <https://doi.org/10.1016/j.radonc.2020.09.041>.
21. Jin H, Su Z. SU-E-T-225: a comprehensive evaluation of real-time motion tracking of a surface imaging system for lung treatment. *Med Phys*. 2012;39:3755.
22. Zhao H, Williams N, Poppe M, Sarkar V, Wang B, Rassiah-Szegedi P, et al. Comparison of surface guidance and target matching for image-guided accelerated partial breast irradiation (APBI). *Med Phys*. 2019;46:4717–24.
23. Zhao B, Park YK, Gu X, Reynolds R, Timmerman R, Sher DJ. Surface guided motion management in glottic larynx stereotactic body radiation therapy. *Radiother Oncol*. 2020. <https://doi.org/10.1016/j.radonc.2020.08.027>.
24. Bertholet J, Knopf A, Eiben B, McClelland J, Grimwood A, Harris E, et al. Real-time intrafraction motion monitoring in external beam radiotherapy. *Phys Med Biol*. 2019;64:15TR01.
25. Chugh BP, Quirk S, Conroy L, Smith WL. Measurement of time delays in gated radiotherapy for realistic respiratory motions. *Med Phys*. 2014;41:091702.
26. Klein EE, Hanley J, Bayouth J, Yin FF, Simon W, Dresser S, et al. Task Group 142 report: Quality assurance of medical accelerators. *Med Phys*. 2009;36:4197–212.
27. Goharian M, Khan RFH. Measurement of time delay for a prospectively gated CT simulator. *J Med Phys*. 2010;35:123–7.
28. Smith WL, Becker N. Time delays and margins in gated radiotherapy. *J Appl Clin Med Phys*. 2009;10:140–54.
29. Nakamura N, Hatanaka S, Shikama N, Akahane K, Sekiguchi K. Quantification of cold spots caused by geometrical uncertainty in field-in-field techniques for whole breast radiotherapy. *Jpn J Clin Oncol*. 2011;41:1127–31.
30. Bortfeld T, Jiang SB, Rietzel E. Effects of motion on the total dose distribution. *Semin Radiat Oncol*. 2004;14:41–51.
31. Kang H, Yorke ED, Yang J, Chui CS, Rosenzweig KE, Amols HI. Evaluation of tumor motion effects on dose distribution for hypofractionated intensity-modulated radiotherapy of non-small-cell lung cancer. *J Appl Clin Med Phys*. 2010;11:3182.
32. Sawant A, Venkat R, Srivastava V, Carlson D, Povzner S, Cattell H, et al. Management of three-dimensional intrafraction motion through real-time DMLC tracking. *Med Phys*. 2008;35:2050–61.
33. Tacke MB, Nill S, Krauss A, Oelfke U. Real-time tumor tracking: automatic compensation of target motion using the Siemens 160 MLC. *Med Phys*. 2010;37:753–61.
34. Depuydt T, Haas OCL, Verellen D, Erbel S, De Ridder M, Storme G. Geometric accuracy evaluation of the new VERO stereotactic body radiation therapy system. In: UKACC international conference on control. 2010. p. 1–6.
35. Sharp GC, Jiang SB, Shimizu S, Shirato H. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys Med Biol*. 2004;49:425–40.
36. Lee SJ. Respiratory motion estimation with hybrid implementation of extended Kalman filter. *IEEE Trans Ind Electron*. 2012;59:4421–32.
37. Kakar M, Nystrom H, Aarup LR, Notttrup TJ, Olsen DR. Respiratory motion prediction by using the adaptive neuro fuzzy inference system (ANFIS). *Phys Med Biol*. 2005;50:4721–8.
38. Murphy MJ, Dieterich S. Comparative performance of linear and nonlinear neural networks to predict irregular breathing. *Phys Med Biol*. 2006;51:5903–14.
39. Narendra KS, Parthasarathy K. Identification and control of dynamical systems using neural networks. *IEEE Trans Neural Netw*. 1990;1:4–27.
40. Cireşan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification. In: *Computer vision & pattern recognition*. 2012.
41. Ernst F, Schlaefer A, Schweikard A. Prediction of respiratory motion with wavelet-based multiscale autoregression. In: *International conference on medical image computing and computer-assisted intervention*. 2007.
42. Ernst F, Schweikard A. Forecasting respiratory motion with accurate online support vector regression (SVRpred). *Int J Comput Assist Radiol Surg*. 2009;4:439–47.
43. Kanoulas E, Aslam JA, Sharp GC, Berbeco RI, Nishioka S, Shirato H, et al. Derivation of the tumor position from external respiratory surrogates with periodical updating of the internal/external correlation. *Phys Med Biol*. 2007;52:5443–56.
44. Schweikard A. Robotic motion compensation for respiratory movement during radiosurgery. *Comput Aided Surg*. 2015;5:263–77.
45. Geneser SE, Hinkle JD, Kirby RM, Wang B, Salter B, Joshi S. Quantifying variability in radiation dose due to respiratory-induced tumor motion. *Med Image Anal*. 2011;15:640–9.
46. Blackall JM, Ahmad S, Miquel ME, McClelland JR, Landau DB, Hawkes DJ. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol*. 2006;51:4147–69.
47. 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;7:483–92.
48. Isaksson M, Jaldén J, Murphy MJ. On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med Phys*. 2005;32:3801–9.
49. Seregini 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;29:48–59.
50. Ernst F, Martens V, Schlichting S, Beširević A, et al. Correlating chest surface motion to motion of the liver using  $\epsilon$ -SVR—a Porcine study. In: *International conference on medical image computing and computer-assisted intervention*. 2009.
51. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9:1735–80.
52. Bradbury J. Exploring two extensions to LSTM machine translation. 2015.
53. Han S, Kang J, Mao H, Hu Y, Li X, Li Y, et al. ESE: efficient speech recognition engine with compressed LSTM on FPGA. 2016.
54. Breerton RG, Lloyd GR. Support vector machines for classification and regression. *Analyst*. 2010;135:230–67.
55. Poels K, Depuydt T, Verellen D, Gevaert T, Dhont J, Duchateau M, et al. Improving the intra-fraction update efficiency of a correlation model used for internal motion estimation during real-time tumor tracking for SBRT patients: fast update or no update? *Radiother Oncol*. 2014;112:352–9.
56. McClelland JR, Hawkes DJ, Schaeffter T, King AP. Respiratory motion models: a review. *Med Image Anal*. 2013;17:19–42.
57. Ernst F. *Compensating for quasi-periodic motion in robotic radiosurgery*. Berlin: Springer; 2012. <https://doi.org/10.1007/978-1-4614-1912-9>.
58. Kuglin CD. The phase correlation image alignment method. In: *Proc Intl Conf Cybernetics and Society*. 1975.
59. Gottesfeld BL. A survey of image registration techniques. *ACM Comput Surv*. 1992;24:325–76.
60. Deutsch. *Supervised sequence labelling with recurrent neural networks*. Berlin: Springer; 2012.
61. Smola AJ, Scholkopf B. A tutorial on support vector regression. *Stat Comput*. 2004;14:199–222.
62. Murphy MJ, Isaksson M, Jaldén J. Adaptive filtering to predict lung tumor motion during free breathing. Berlin: Springer; 2002.
63. Yang J, Tan KK, Santamouris M, Lee SE. Building energy consumption raw data forecasting using data cleaning and deep recurrent neural networks. *Buildings*. 2019;9:204.
64. Jiawei X, Murata T. Stock market trend prediction with sentiment analysis based on LSTM neural network. In: *International multiconference of engineers and computer scientists*. 2019. p. 475–9.

65. 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;6:51262–8.
66. Guo J, Xie Z, Qin Y, Jia L, Wang Y. Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM. *IEEE Access*. 2019;7:42946–55.
67. Yessenbayev Z, Makazhanov A. Character-based feature extraction with LSTM networks for POS-tagging task. In: 2016 IEEE 10th international conference on application of information and communication technologies (AICT). 2016.
68. Zhang H, Zhang Q, Shao S, Niu T, Yang X. Attention-based LSTM network for rotatory machine remaining useful life prediction. *IEEE Access*. 2020;8:132188–99.
69. Shao H, Soong B. Traffic flow prediction with Long Short-Term Memory Networks (LSTMs). In: IEEE region 10 conference. 2016. p. 2986–9.
70. Luo T. Research on decision-making of complex venture capital based on financial big data platform. *Complexity*. 2018;2018:1–12.
71. Thapa S, Zhao Z, Li B, Lu L, Fu D, Shi X, et al. Snowmelt-driven streamflow prediction using machine learning techniques (LSTM, NARX, GPR, and SVR). *Water*. 2020;12:1734.
72. Park JC, Park SH, Kim JH, Yoon SM, Song SY, Liu Z, et al. Liver motion during cone beam computed tomography guided stereotactic body radiation therapy. *Med Phys*. 2012;39:6431–42.
73. Murphy MJ, Isaakson M, Jalden J. Adaptive filtering to predict lung tumor motion during free breathing. In: 16th international congress and exhibition on computer assisted radiology. Paris, France: Springer; 2002. p. 539–44.
74. Ernst F, Schlaefer A, Schweikard A. Predicting the outcome of respiratory motion prediction. *Med Phys*. 2011;38:5569–81.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

